An abstract illustration on the left side of the slide. It features stylized, colorful silhouettes of two people in profile, facing each other. They appear to be interacting with a large, central digital screen or interface. The interface has a green checkmark icon on a blue background, suggesting a ballot or a vote. The overall style is graphic and modern, using bold colors like orange, teal, and yellow.

Digital Democracy & AI Alignment

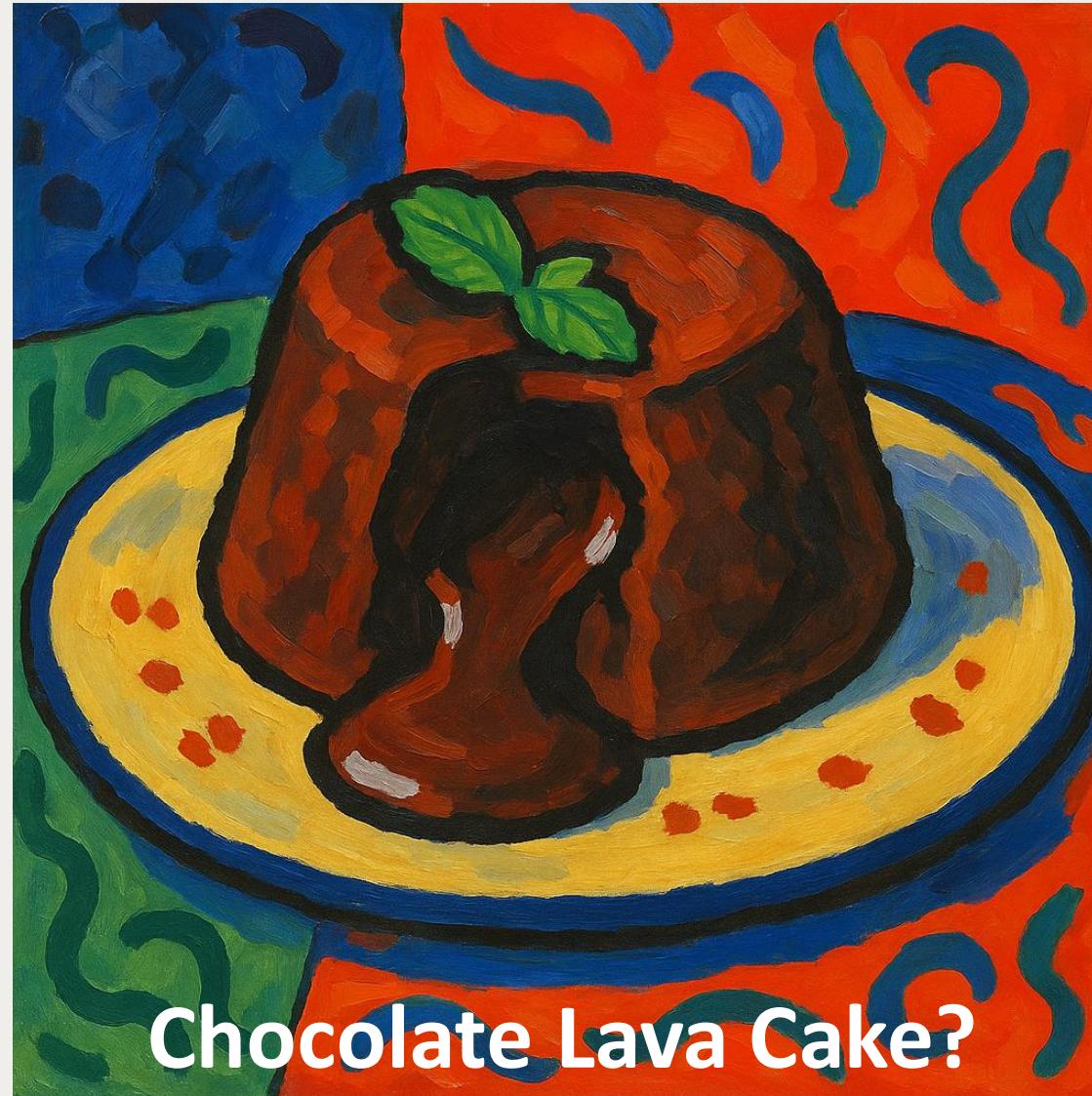
23.01.2026
Uni. Konstanz

Joshua C. Yang
Computational Social Science
ETH Zurich

The University of Konstanz has decided to provide an unlimited supply of one dessert for students to enjoy year-round. Which dessert should it be?

1. Banana Walnut Muffin
2. Strawberry Cream Tart
3. Blueberry Crumble Muffin
4. Chocolate Lava Cake
5. Pear Tart

The winner is...



Chocolate Lava Cake?

Context: Democracy is not always so fair

2000 US election in Florida

Party	Candidates	Votes	Electoral votes
Rebuplican	George W. Bush	2,912,790	25 <input checked="" type="checkbox"/>
Democrat	Al Gore	2,912,253	0
Green	Ralph Nader	97,488	0

Spoiler effect: Because of a third candidate, Bush won by just **537 votes** in Florida to win the whole presidency.



Duverger's Law

Core idea

Plurality voting in single member districts usually leads to a two party system. The mechanism is both mechanical and psychological.

- **Mechanical effect**

Only the candidate with the most votes wins the seat. Smaller parties rarely gain representation, which discourages their continued participation.

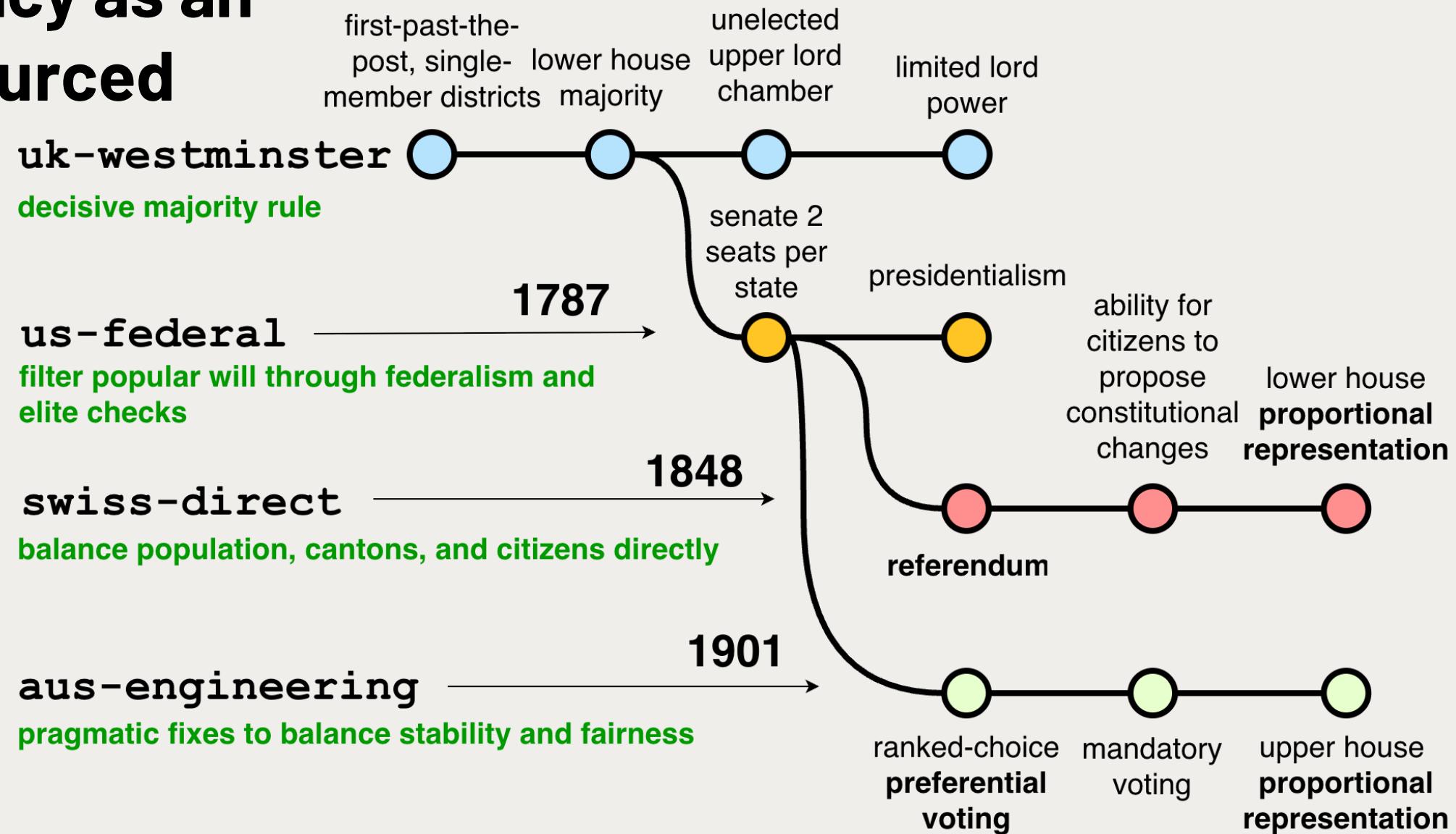
- **Psychological effect**

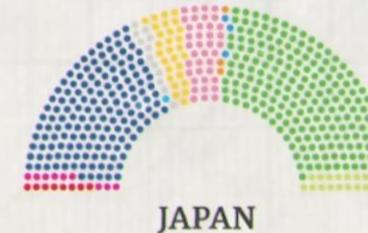
Voters tend to avoid "wasting" their vote on a candidate who is unlikely to win. They strategically choose the major party they dislike least. Politicians also strategically join larger parties instead of building small ones.

Result

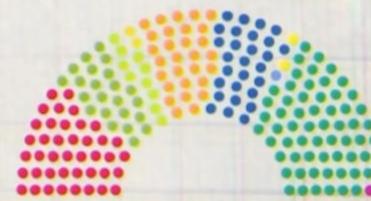
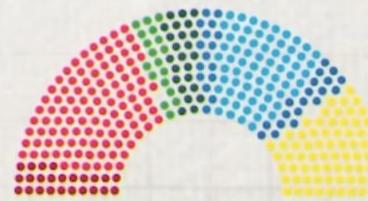
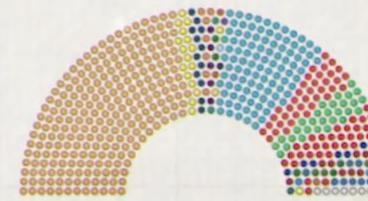
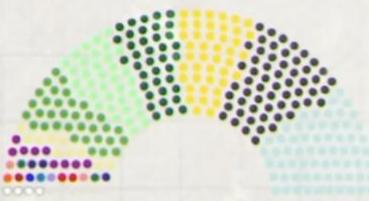
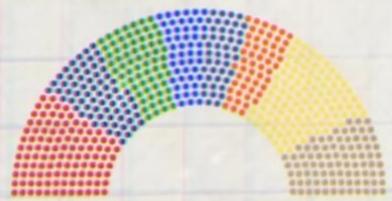
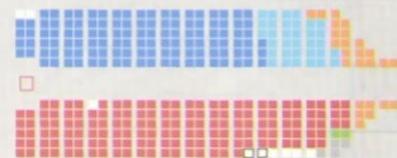
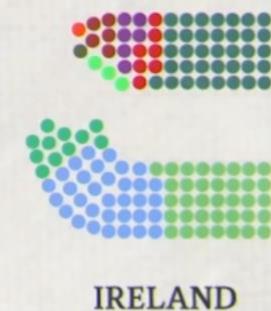
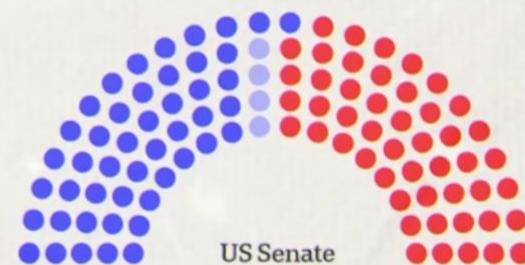
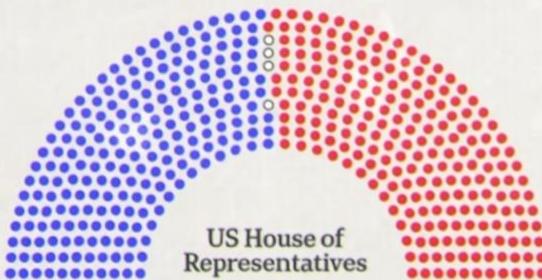
Over time, political competition converges toward two dominant parties.

Democracy as an Open-Sourced Project





Our voting system doesn't just count the votes— it shapes the choices and the society.





French Revolution and Social Choice

French Revolution: Transitioning from “Divine Right” to “Reason”:

If the new government was to be based on the **“Will of the People”** (Rousseau's concept of the *General Will*), how do you scientifically measure and enact that will from millions of diverse, often conflicting, preferences?

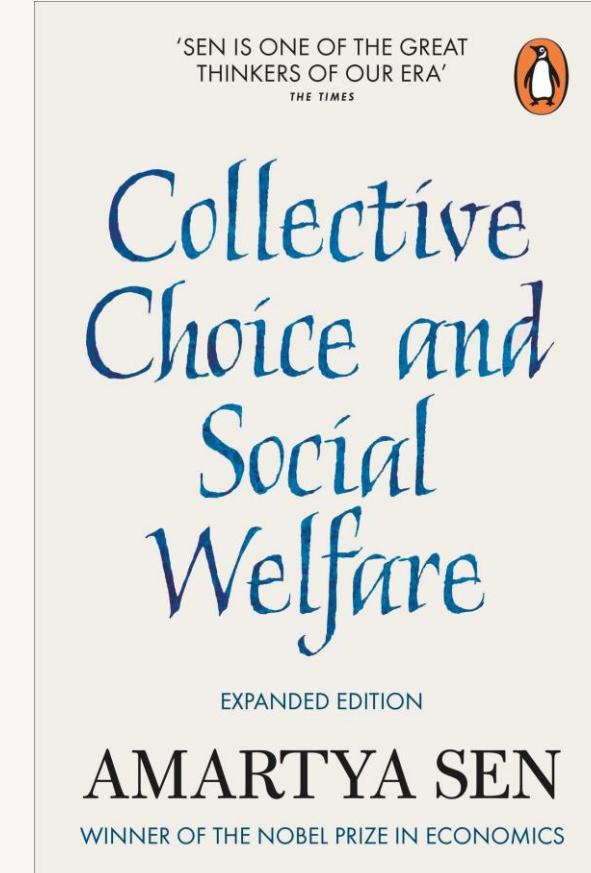
Background: The Shift in Social Choice

**From: "Democracy is
mathematically impossible, so
pick the least-bad voting method"**

e.g. Arrow's Impossibility Theorem

**To: "Design systems that gather
richer information about citizen
preferences and values"**

e.g. Sen's Collective Choice and Social Welfare



Context: Digitalisation enables new participation

The Helsinki of dreams is created together. OmaStadi.

OmaStadi is Helsinki's way of implementing participatory budgeting. In OmaStadi, you can submit proposals and, through voting, decide how the city uses €10 million. A new round starts in the autumn of 2025, and you can submit proposals on this website from September 22nd to October 5th, 2025. The city will implement the proposals that receive the most votes.

...



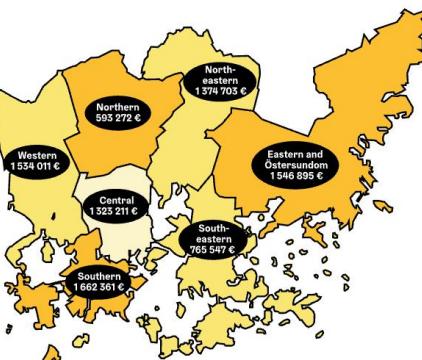
1 2 3 4

02.10.2023 - 15.10.2023 Proposals
28.02.2024 - 18.03.2024 Voting
08.04.2024 - 21.04.2024 Co-creation
From autumn 2024 Implementation

Project planning and implementation is underway.

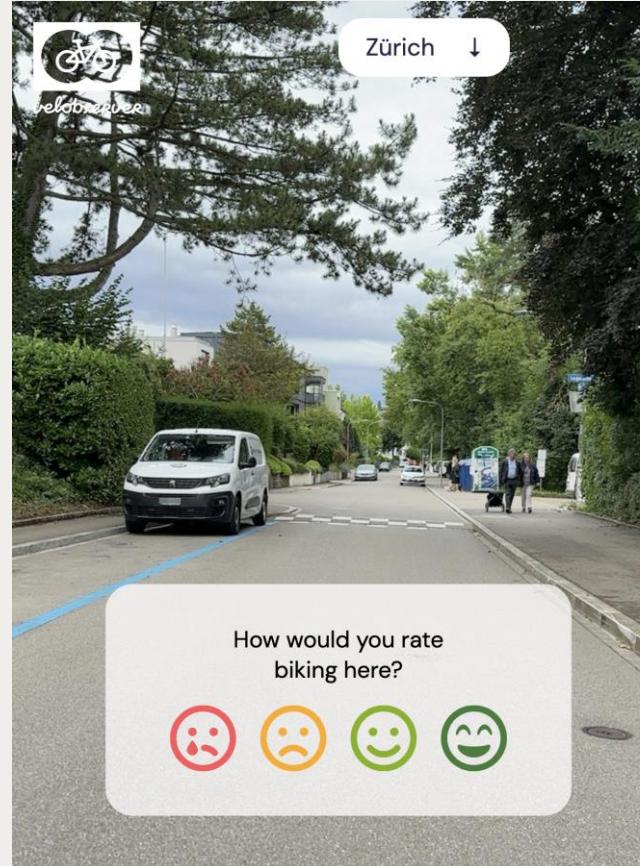
Implementation of the proposals voted on in 2024 has begun. Check out how and when the proposals will be carried out in your living area.

Check out the OmaStadi-projects →



District	Budget Allocation (€)
Northern	1 374 703
Western	1 534 016
Central	1 525 211
Southern	1 662 361
South-eastern	765 547
Eastern and Östermalm	1 546 895

Participatory Budgeting

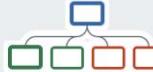


Crowdsourcing

Are open borders a good idea?

Use Template

i



Open borders are a good idea.

Pros + Cons +

Open borders increase economic development. Closed borders protect national Comments and Edit History. Click to open

Open borders promote exchange among people. Closed borders protect cultural heritage.

Deliberative Platforms

Context: Digitalisation enables richer information

Combined approval voting

Approve, disapprove, or abstain for each candidate.

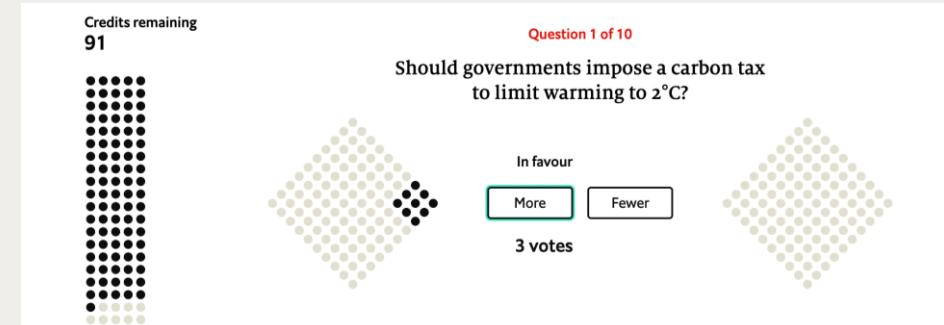


Rank voting

Rank the candidates. The system then assigns a score accordingly.



Quadratic voting



Cumulative voting

You have 10 points in total to distribute.

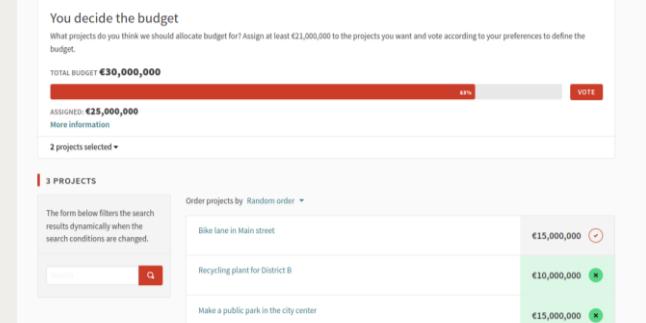


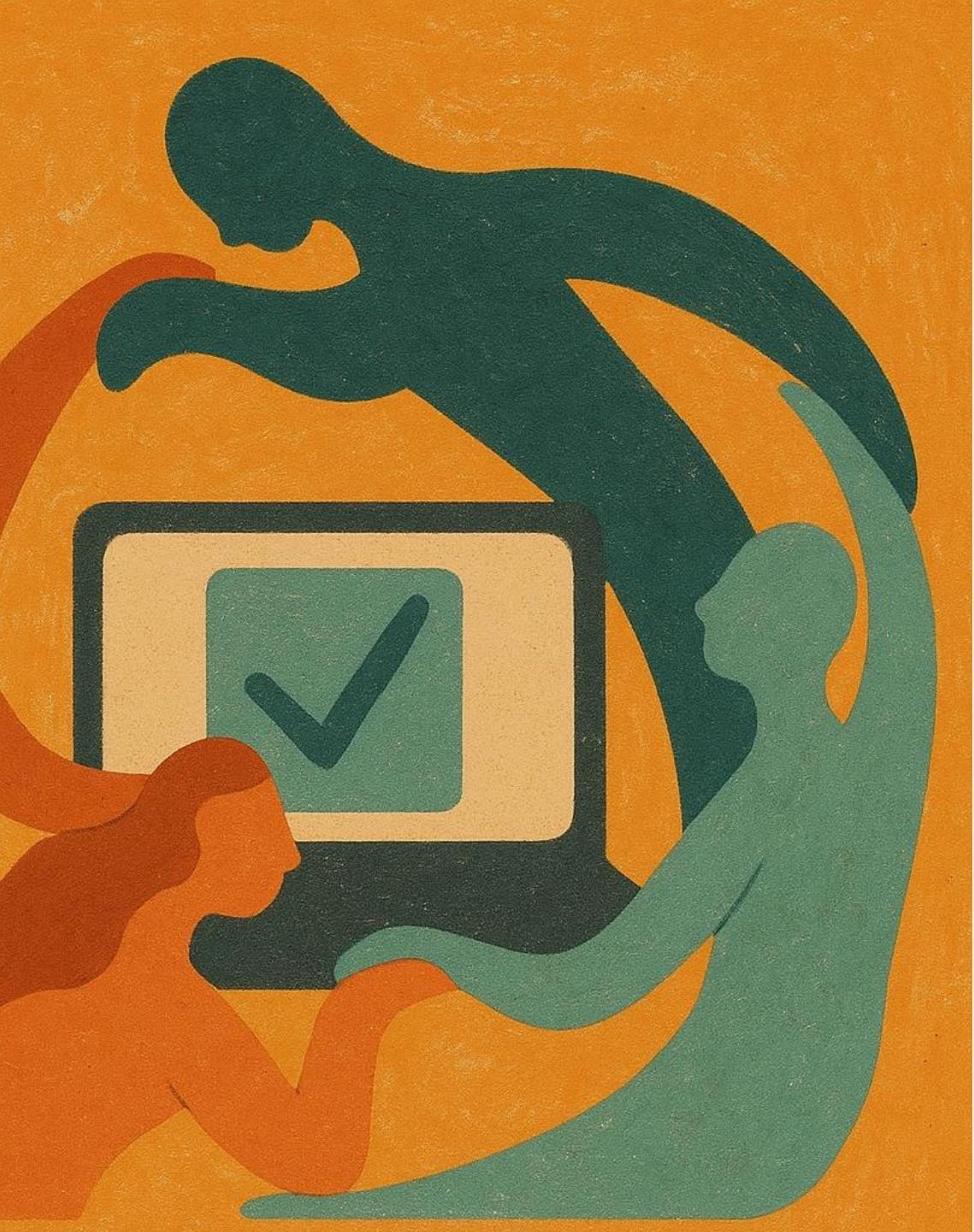
Available points



into the ballot box

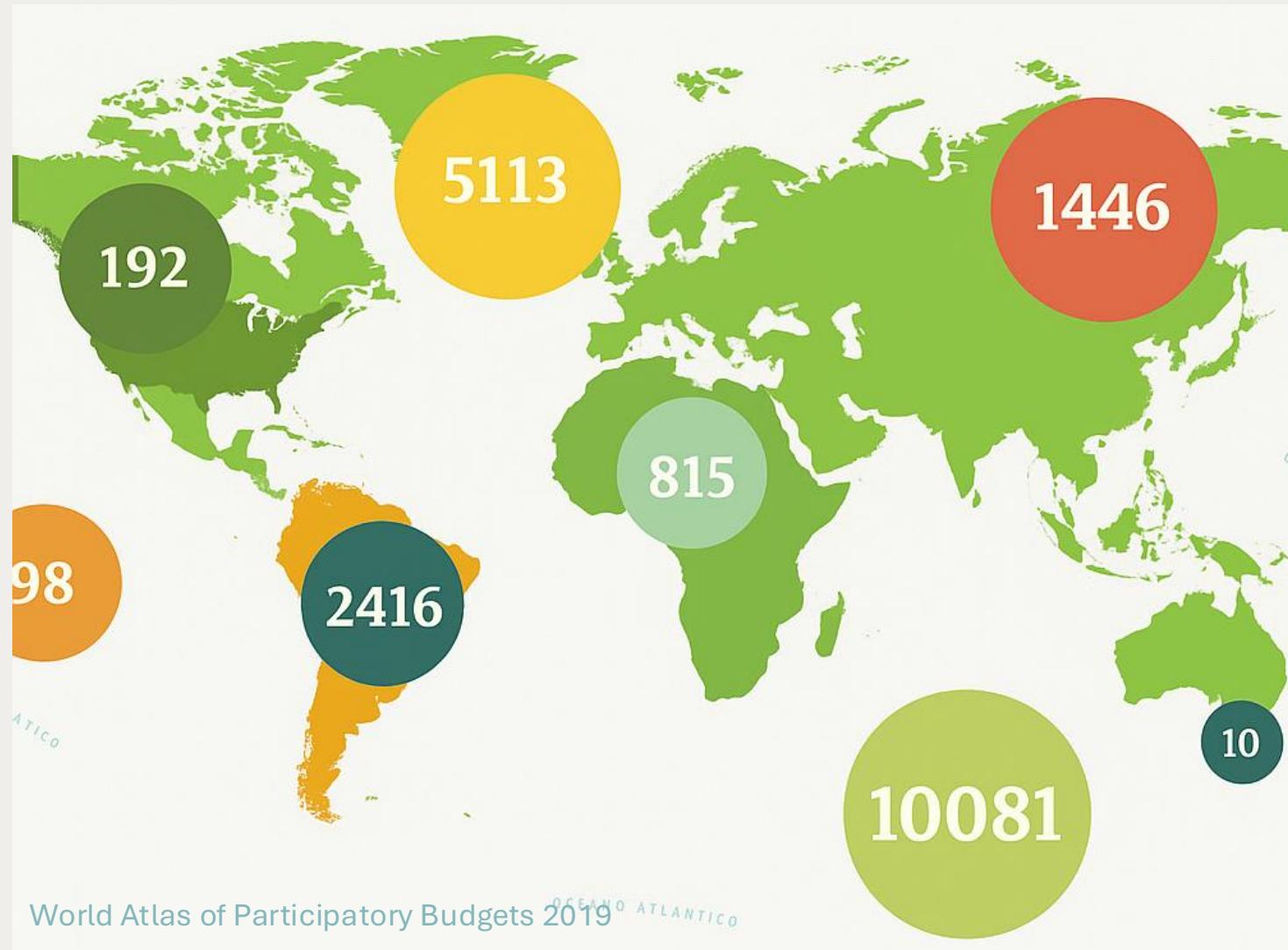
Knapsack voting



An abstract illustration on the left side of the slide. It features stylized, flowing shapes in dark teal, orange, and light beige against a bright orange background. In the center, there is a graphic element resembling a smartphone or a computer monitor displaying a large green checkmark inside a white square with a black border.

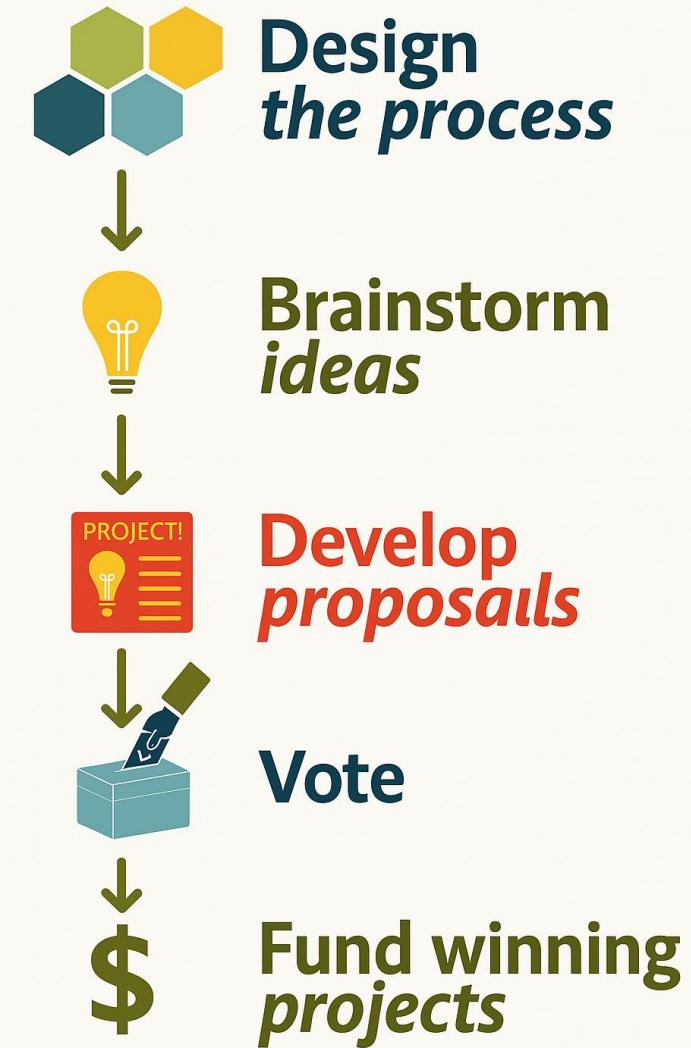
Participatory Budgeting

Background: Participatory Budgeting



Background: Participatory Budgeting

Paris: 100 million per year and 5% of the city budget.



Background: Participatory Budgeting



Overview

Ideas

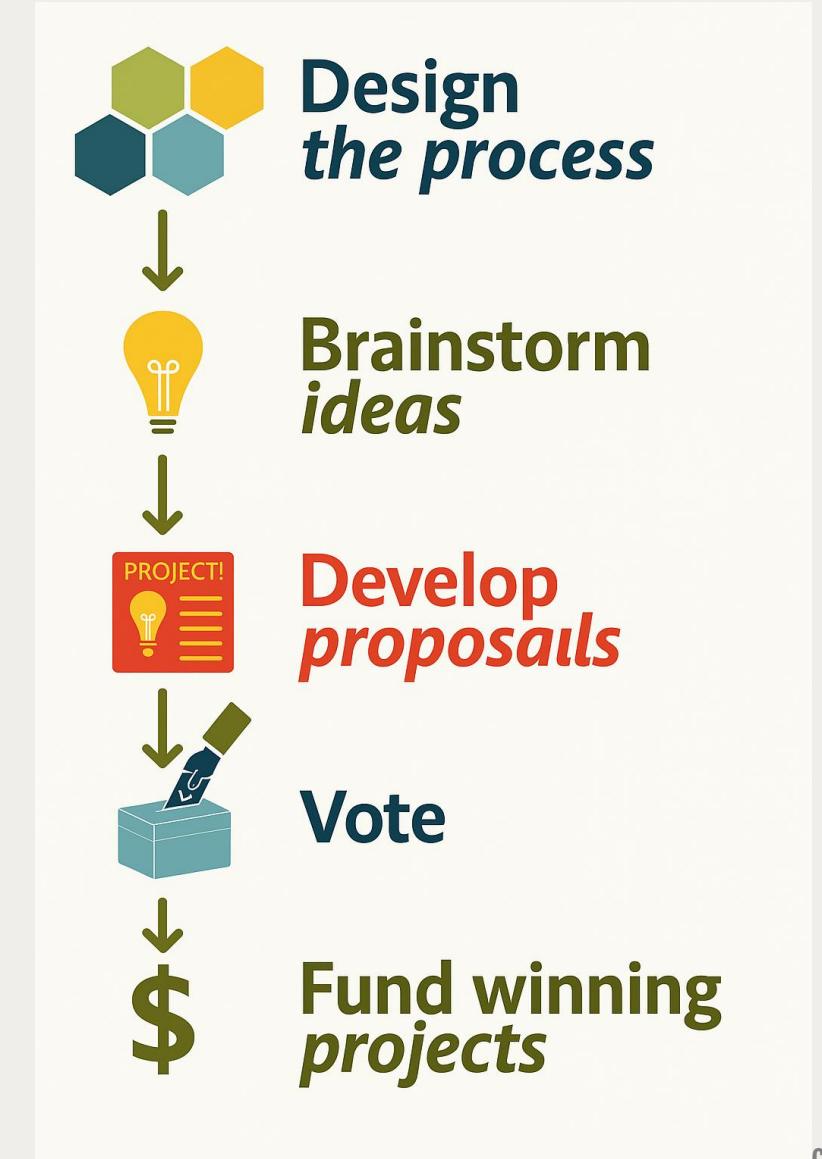
vote

Evaluation

Implementation

(I A A A)

Category	Image	Description	Status
Implementation		Open gymnasium at the Aare school: games and fun on Saturday evenings	Umgesetzt!
Ideas		Public bicycle pumps	nicht umgesetzt!
Implementation		A hedge creates usable space	nicht umgesetzt!
Implementation		Let's Play Football in Scheibenchachan	Umgesetzt!
Implementation		Boule for everyone in the Telli	Umgesetzt!
Ideas		Summer fun in the Sonnmatt summer garden	Umgesetzt!
Ideas		Garden for All	Umgesetzt!
Implementation		Public herb garden	Umgesetzt!



Votes i

August 2025				
	Thu 28		Fri 29	
	10h-12h	13h-15h	15h-18h	10h-12h
Rafael	✓	✓	✓	✓
Joshua	✓	✓	✓	✓
Philip				(✓)
Jonas				
Josh Y.	✓	✓	✓	✓
Mateusz				
Total	3	3	3	3
6 polled users				(+1)

Decisions can be made better in an easy way

Fondue Democracy 🥤 :

by Josh · 21 hours ago

★ Email Ambush: Agnieszka sends an email and the first N people to answer get the seats. 39 points

✉️ Human Twister: players follow balance commands and the first four who fall are out. 37 points

🎲 Python random: draw a random.sample() selects names at random from the list. 36 points

🍺 Knock-out Test: drink together and the first four who cannot stay upright lose their seat. 34 points

👤 Merit Based: seats go first to the people who have stayed in COSS longer 30 points

📦 Everyone draws lots: draw from a pool of hidden tickets until all "yes" tickets and four "no" tickets are revealed. 29 points

Which destination should host the COSS Retreat 2026? (7 pairs would be enough)

Hotel Meielisalp, Leissigen (Lake Thun) – Mountain views, panorama trail hikes, and BBQ evenings.



or

Casa San Bernardo, Contra (Lake Maggiore) – Hills above Locarno with hiking, climbing, and lake trips.



Background: Bounded Discrete Participatory Budgeting

Voters



Available Budget



CHF 100

Bounded Budget: A upper limit on the total cost of selected projects

Project Cost



CHF 20



CHF 20



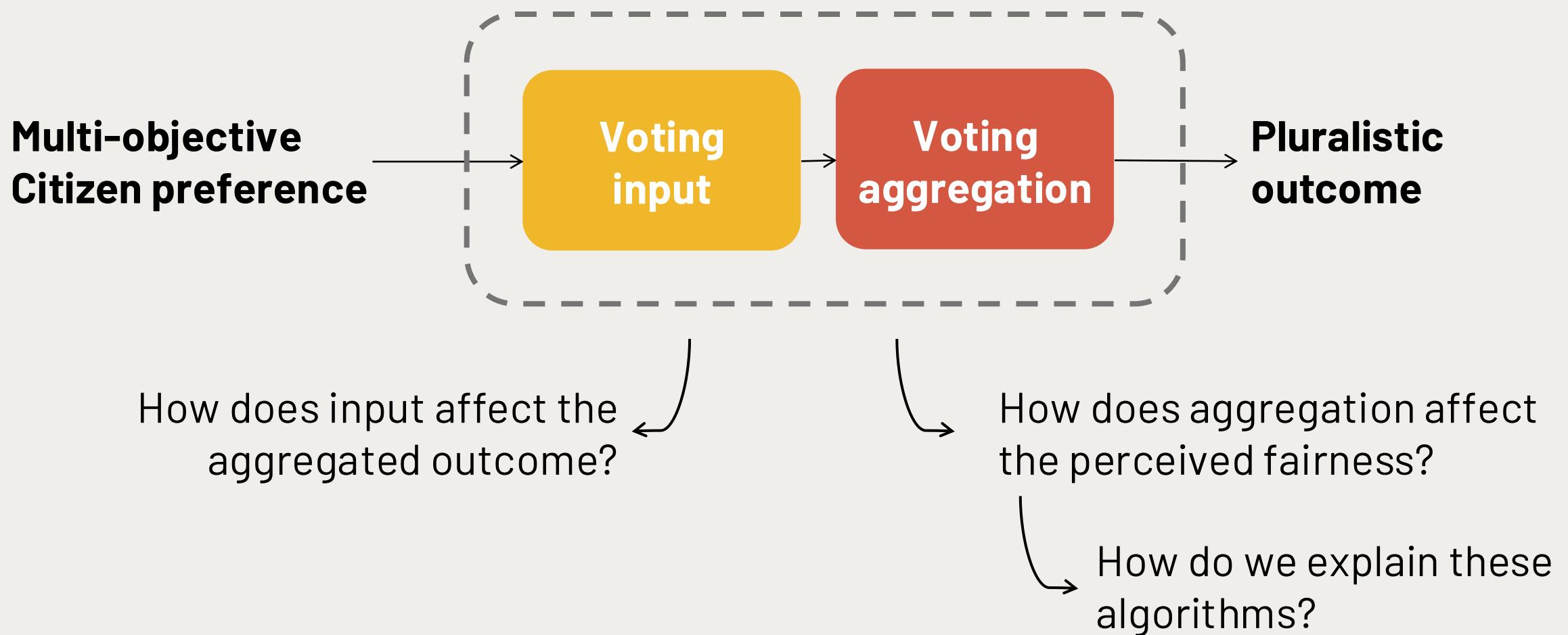
CHF 30



CHF 40

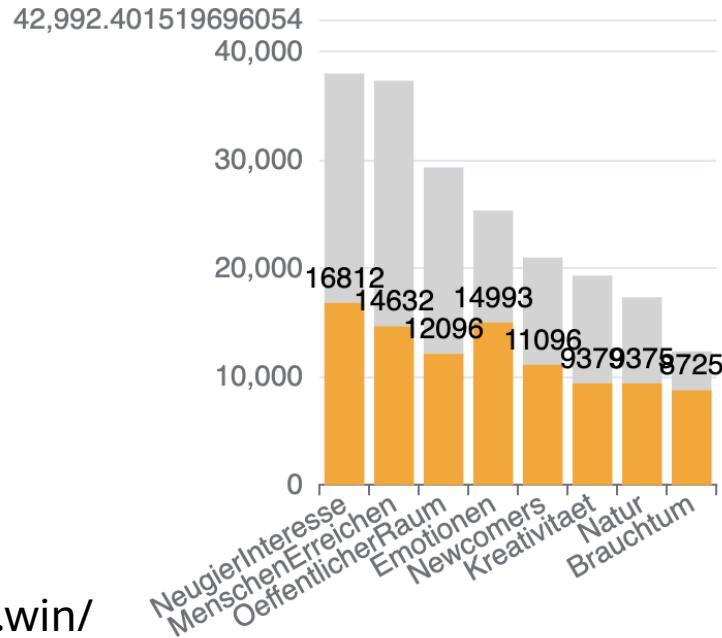
Discrete Projects: Projects can only be funded in full or not at all.

Background: Voting input and aggregation



Priceability as Explainability

Gruppenbudgetnutzung



<https://kk25.win/>

Budgetübersicht

Gruppen: 189,732
(50% des gesamten Budgets ausgegeben) **Einzelpersonen: 1**
(50% des gesamten Budget)

Insgesamt ausgegeben: 378,901
(95% des Gesamtbudgets)

🔍 Neugier		TOTAL
		37,164
ITEM	VOTES	SPENT
Vélo volant – Das Figur...	9	3610
Tier & Text	4	4544
1 Jahr Tüfteln in der Fa...	4	3262
"Silvesterabend" – eine...	3	8435
Winterthurer Druckwoc...	3	1646
Das Winterthurer Publi...	3	0
Die Immigranten / The ...	3	0
Kinder- sowie Jugendt...	2	4684
LES FEMMES FATALES	2	2324
Samstagsmatinée	1	2620
Rauk-it	1	2582
Der Wurm in meinem K...	1	1929
Never Enough	1	1051
Tag der Vielfalt 2025	1	478
Subtotal	38	37,165

❤️ Emotionen		TOTAL
		21,243
ITEM	VOTES	SPENT
Nicht schuld!	5	3177
Eisblumen – ein etwas ...	5	1939
Tuuli	4	5231
Vélo volant – Das Figur...	3	1173
Green Rock	2	2969
Schreibprojekt «Palazz...	2	2866
LES FEMMES FATALES	2	2326
Never Enough	1	1056
Konzert Kammerchor u...	1	506
Ruhe (un)sanft!	1	0
Subtotal	26	21,243

🌿 Natur		TOTAL
		9,371
ITEM	VOTES	SPENT
ANTE Konzertreihe	4	1297
Winterthurer Wolfs-Qu...	3	3564
Tier & Text	3	3351
Vélo volant – Das Figur...	3	1159
RaumzeiTraum	3	0
„Poems of Grief“ – „Ge...	1	0
Iauschig: Literatur und ...	1	0
Subtotal	18	9,371

🌿 Newcomers		TOTAL
		27,339
ITEM	VOTES	SPENT
Open Art Festival	5	4099
Ist das Kunst oder kan...	4	8816
Liebesfiguren und kom...	3	4425
Alli Vibes Geile Scheiss	3	3333
Plattform nachhaltiges...	2	3061
Erzählladen auf Zeit	2	1636
Ausgetriggert -wenn Tr...	2	705
Vielfalt im MFW-Moder...	1	1265
Subtotal	22	27,340

Voting aggregation: Greedy Method

Project Name	Cost €	Votes	Greedy
A green wall at the intersection of Blancs Manteaux and Archives streets	30,000	788	✓
A community café in the Tour Saint-Jacques neighbourhood	15,000	706	✓
An arts room for the Saint-Merri Renard school	300,000	702	✓
Model energy renovation of a school in the 4th arrondissement	1,000,000	655	✓
Greening of Arsenal Street	120,000	649	✓
A Charlemagne middle school accessible to people with reduced mobility	200,000	630	
Making the entrance hall of the Saint-Merri swimming pool a welcoming space	100,000	528	
Fitness equipment at Place des Vosges	15,000	491	
Showcasing the stones of the Bastille prison	20,000	473	
An electric wheelchair to break isolation	5,000	453	
Creating a canopy over a playground in a square in the 4th arrondissement	150,000	410	
“The Elevator” a third space for equal opportunity open to the world	350,000	315	
Signposting the Compostela pilgrimage route in the 4th arrondissement	10,000	265	
Bike racks on Reynie Street	10,000	240	
Improving access to the Ourscamp cellar	120,000	228	

Paris PB 2019

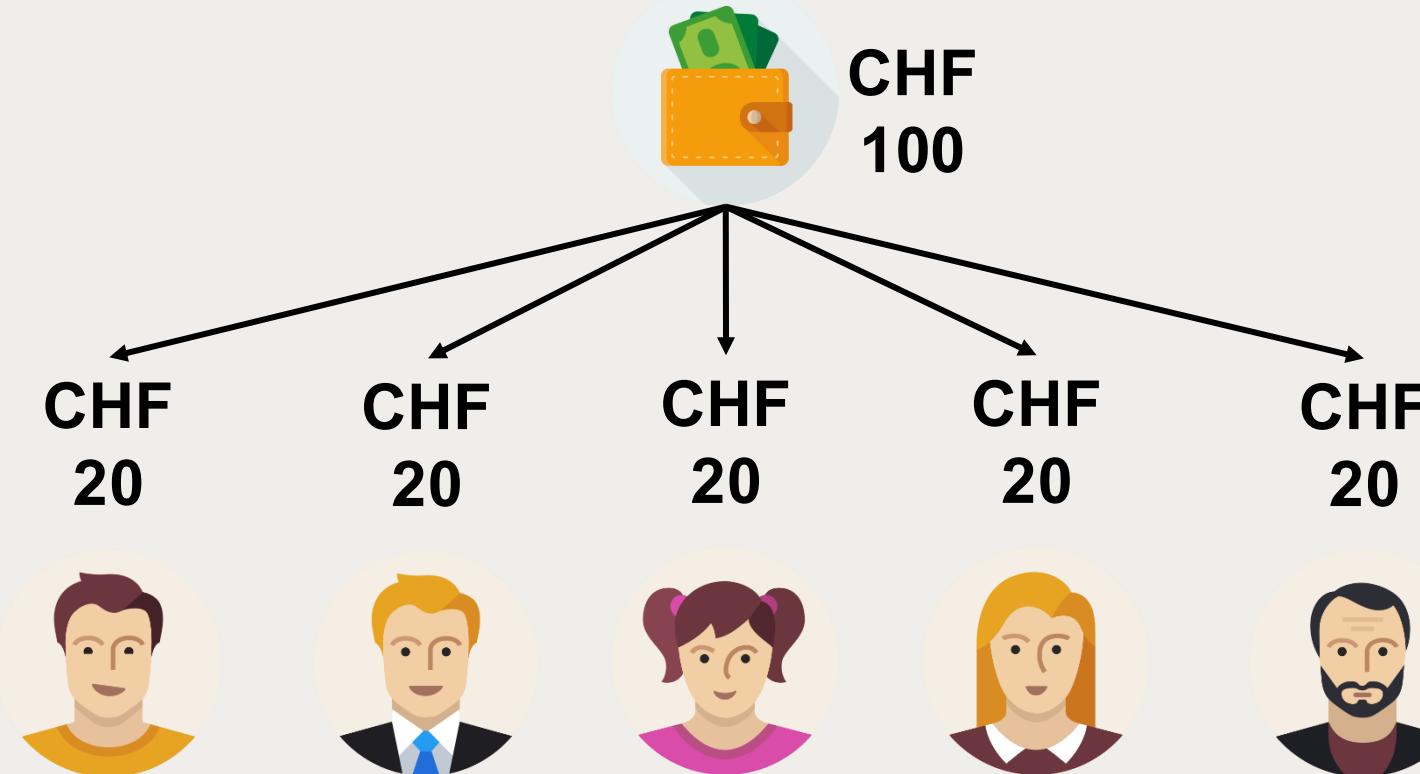
Total budget:
€1,500,000

Problem: some voters can be getting a disproportional amount of budget

Background: Method of Equal Shares (MES)

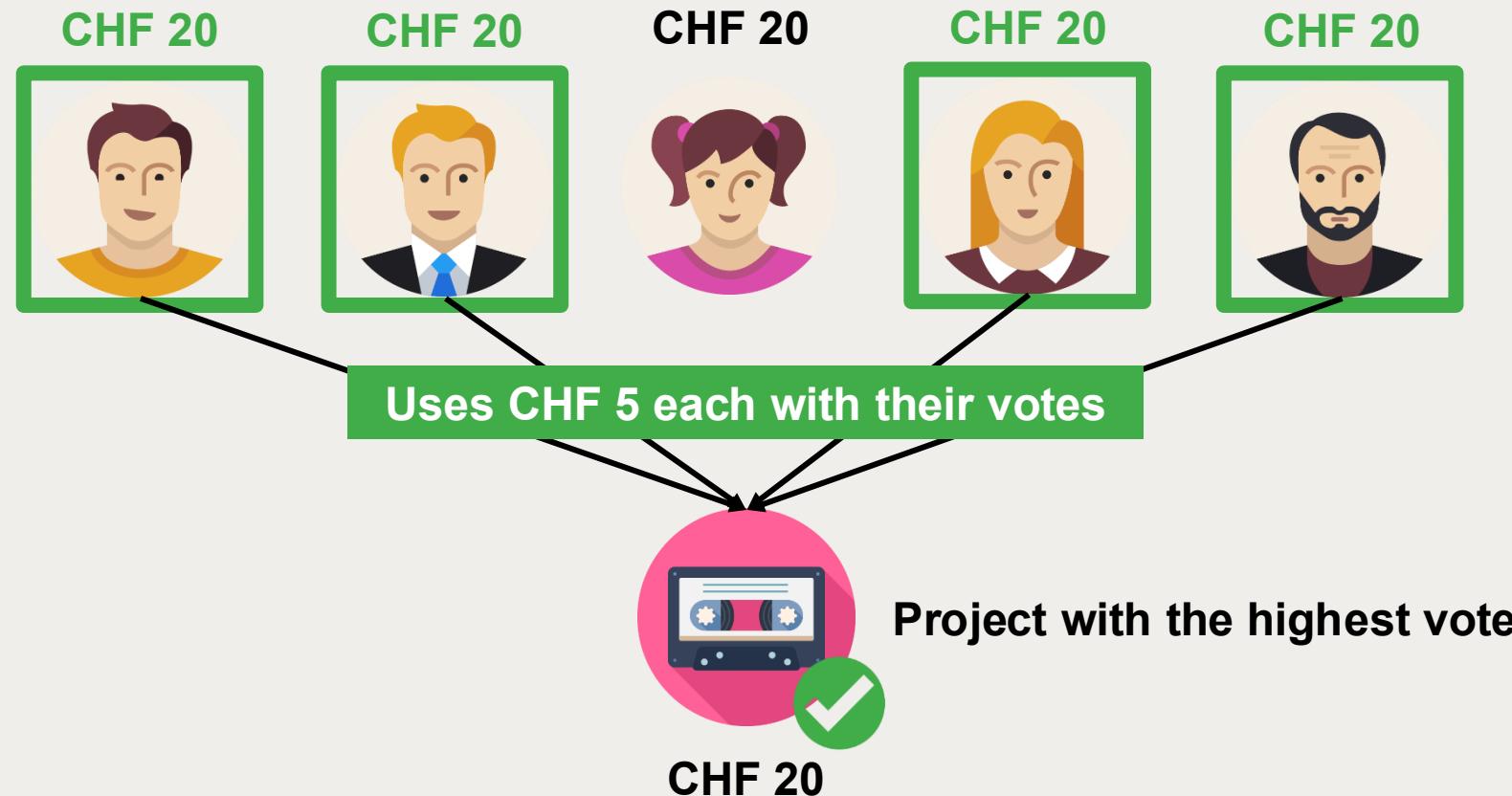
Each voter $i \in N$ receives an **equal starting budget** $b_i \geq 0$.

Let $u_i(p)$ be the utility that voter i assigns to project p . Let c_p be the cost of project p .

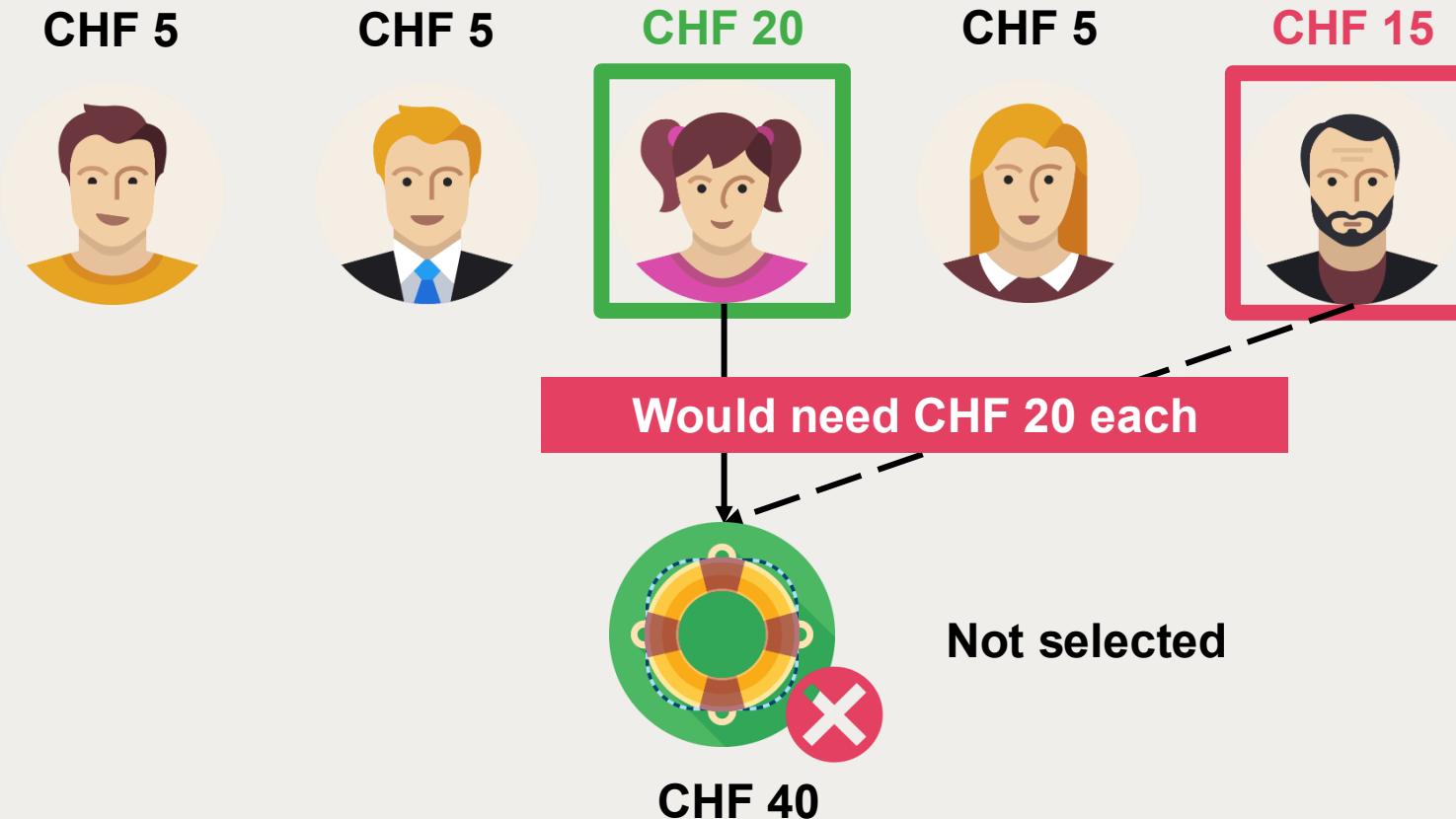


*Peters, Dominik., Pierczyński, G., & Skowron, P. (2021). **Proportional participatory budgeting with additive utilities.** *Advances in Neural Information Processing Systems*

Background: Method of Equal Shares (MES)*



Background: Method of Equal Shares (MES)*



Background: Method of Equal Shares (MES)*

- Proportional to what the voters want
- But the results are not easy to understand

	Project	Cost	Votes	Selected by MES
1	Pfasyl Aargau	3'600	1'455	Yes
2	Sommerliche Kinonächte in der Badi	10'000	1'239	Yes
3	Wildbienen	20'000	1'203	No
4	Öffentliche Velopumpen	4'000	1'191	Yes
5	Kufa - Kulturfest Aarau	15'000	966	No
6	Offenes Kinderatelier	10'000	862	No
7	Ort für alle (F. Fröhlich und H. Hässig)	17'000	817	No
8	Eine Hecke schafft nutzbaren Raum	1'000	800	Yes

Human experiment

Study Setup

- **Formal Registration:**
 - Pre-registered on AEA ACT Registry
 - Approved by ETH Zurich Ethics Commission (approval number: EK 2022-N-143)
- **Platform:** Conducted online via Qualtrics in March 2023

Participants

- **Sample:** 180 Zurich university students
- **Engagement:** High engagement with many participants providing detailed responses in free-text fields

C. Spread 5 points

Now, vote again in the same program. In this vote, you distribute 5 points among the projects you like. You can concentrate your points on one project, or spread them across a few projects.

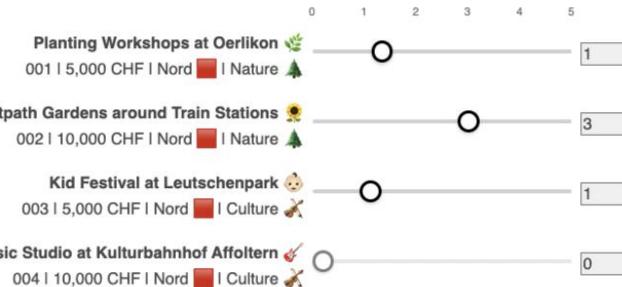


Fig. 20. Screenshot of the D5 Method Interface

E. Select 5 and rank

Now, vote again in the same program. In this vote, you first drag the top 5 projects that you like to the box, then rank them in order. The top project will get 5 points; the second project, 4 points; the third, 3 points; and so on.

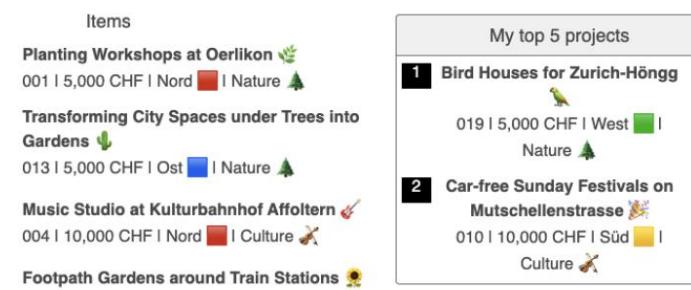


Fig. 22. Screenshot of the S5R Method Interface

Human experiment

Voting Input

1. Vote on city projects using different voting inputs
2. Rate their experience on voting inputs

Voting Aggregation

3. Shown different outcomes calculated using different aggregation methods
4. Rate satisfaction, fairness, trustworthiness

Algorithm Explanation

5. Shown different types of explanations
6. Rate satisfaction, fairness, trustworthiness again

Experimental set-up: 6 Voting inputs

Inputs	Category	Related Voting Rule
Select any number of projects	Approval-based	Approval Voting
Select exactly 5 projects	Fixed-k approval	k-Approval
Distribute 5 points across projects	Cardinality (score)	Cumulative Voting (5 points)
Distribute 10 points across projects	Cardinality (score)	Cumulative Voting (10 points)
Select and rank 5 projects	Ordinal (ranked)	Borda / STV / Ranked Choice
Distribute 10 points across 5 selected	Hybrid (approval + cardinality)	Weighted k-Approval / Hybrid

Experimental set-up: 2 Voting aggregations

	Logic	Representation	Note
Greedy Rule	Picks top-voted projects until budget runs out	Winner-takes-all	The default in PB programs
Method of Equal Shares (MES)	Funds projects if supporters can afford them with their share	Proportional	Theoretical and largely untested in real-world contexts

Experimental set-up: Algorithm Explanations

Pre-hoc explanation

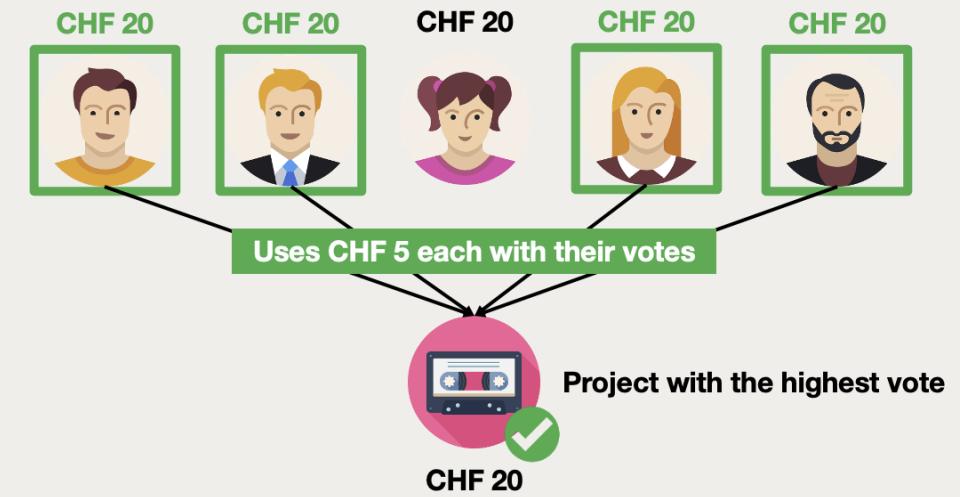
(the **Mechanism** explanation)

This explains *how* the algorithm works. Also called "explanation proper" or "model transparency,"

Post-hoc explanation

(**Individual** and **Group** explanations)

These explain the *outcomes* after the algorithm has run, without necessarily revealing the internal workings



Result: Voting input

- **Expressiveness > ease of use:** people prioritise meaningful representation over convenience in democratic participation.
- **Ranked voting and cumulative voting offer stable outcomes**

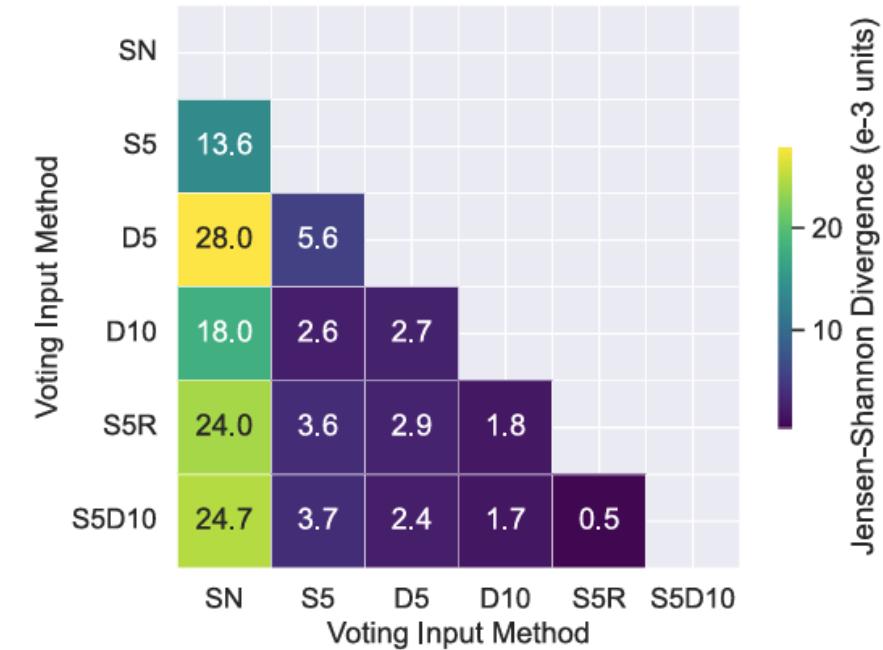
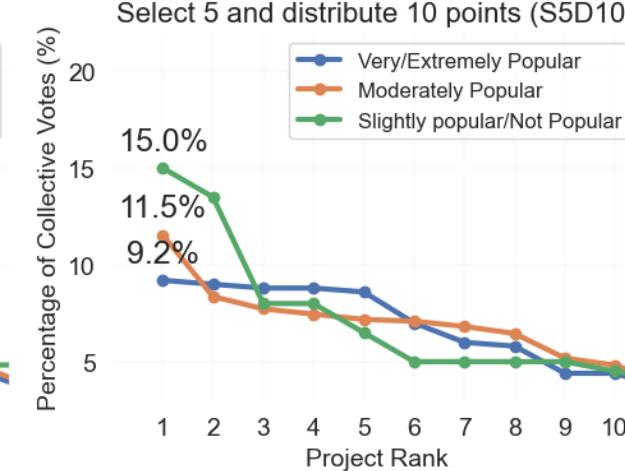
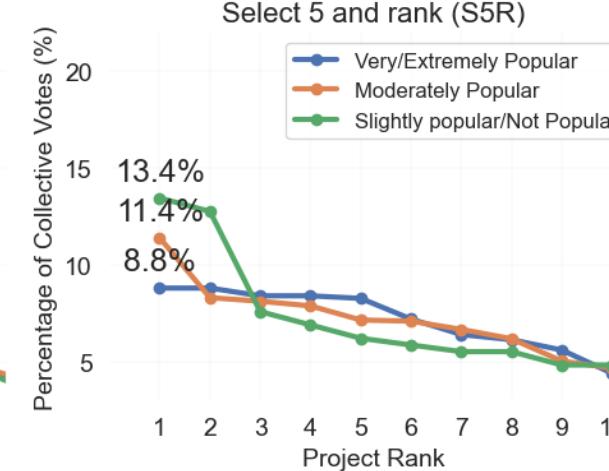
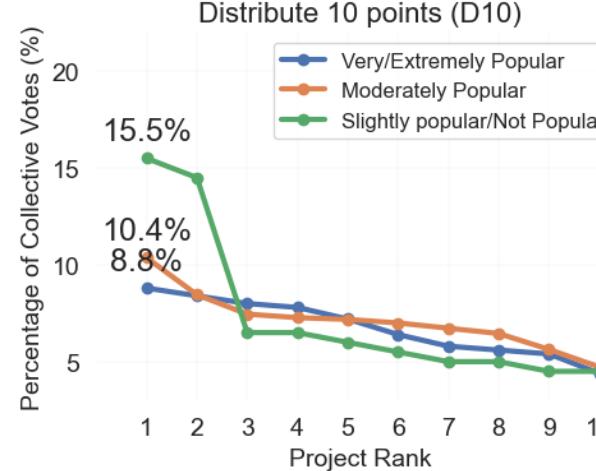
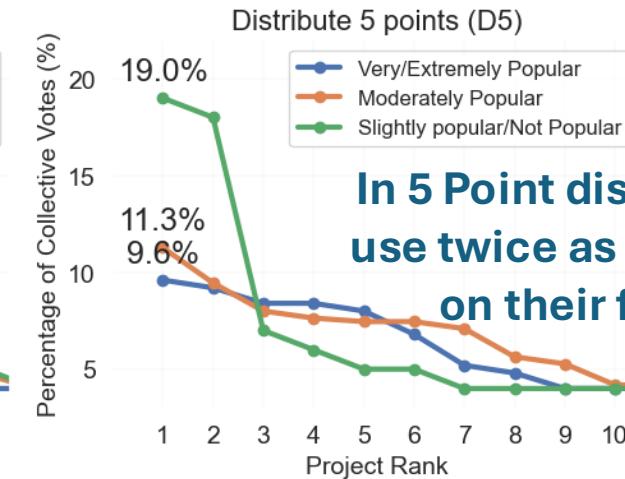
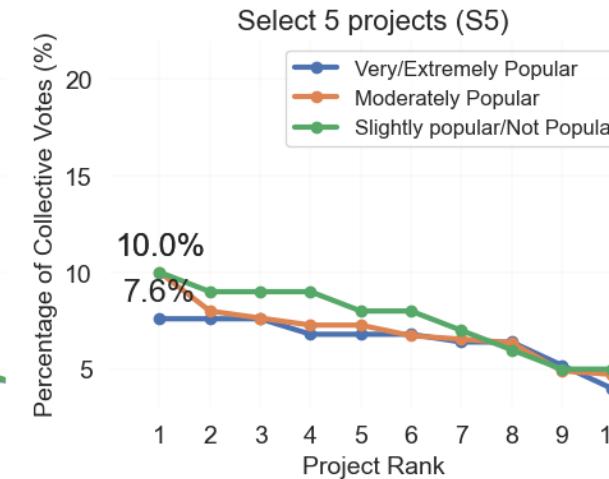
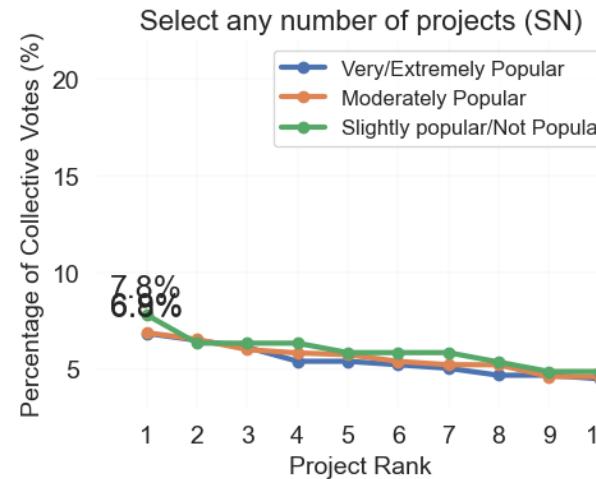


Fig. 4. Heat map of Jensen-Shannon divergence values comparing the distribution of votes between different voting input formats. The higher the value, the larger the difference between the distributions.

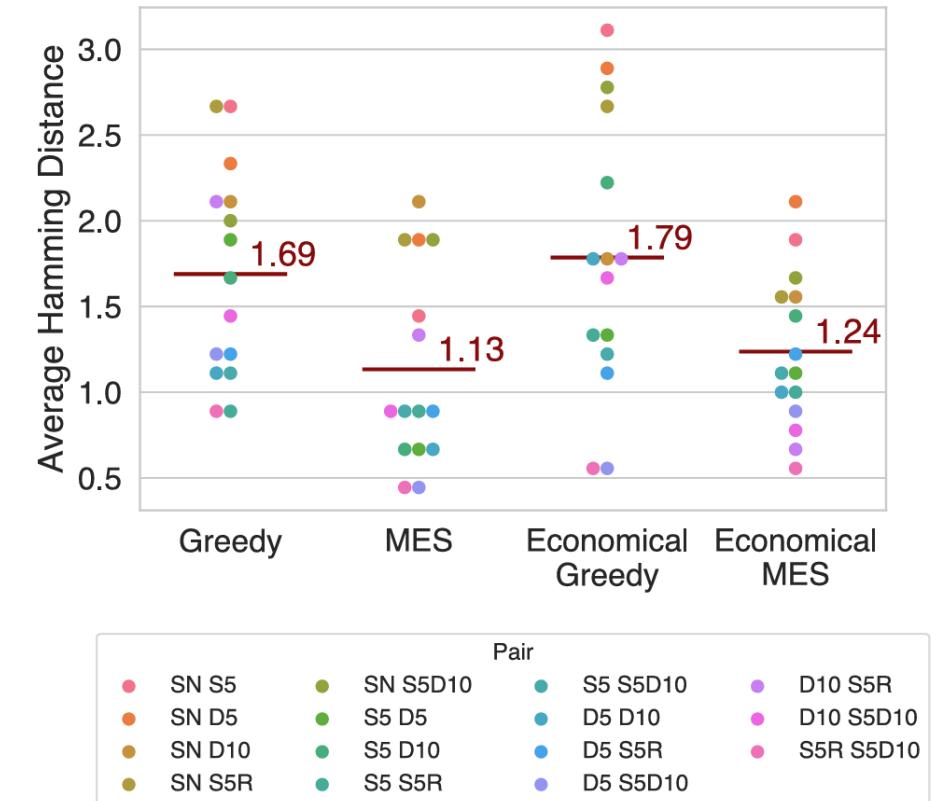
Result: Voting Input

Self-perceived minorities protect their interests in Cumulative Voting

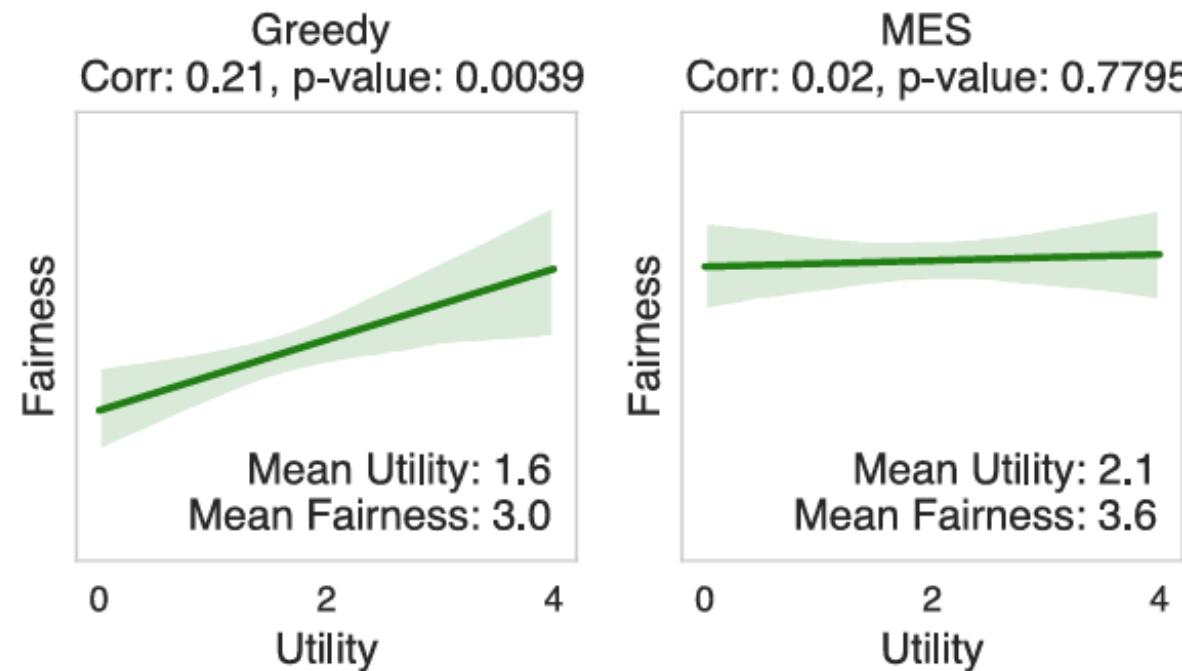


Result: Voting Aggregation

- **MES consistently outperformed Greedy** across all measures:
Satisfaction: 3.44 vs 2.83 (79% rated MES higher), Fairness (after explanation): 3.93 vs 2.97
- **MES is more stable** when voting input formats change: MES: 1.13 average difference / Greedy: 1.69 average difference



Result: Outcome perception



Perceived fairness correlates with winning in simple Greedy

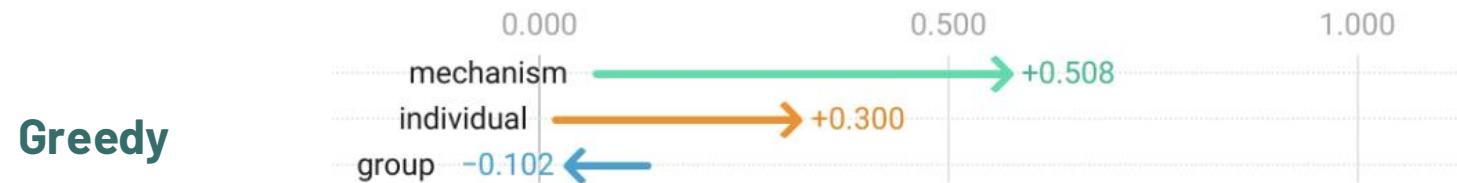
=> **participants finds a simple method fair only when they actually win**

In MES, there is no significant correlation between fairness and utility

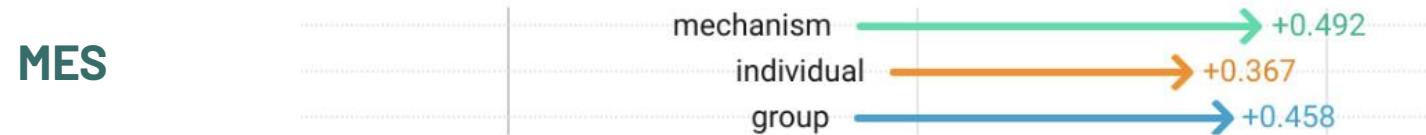
=> **participants find MES fair even if they lose**

Result: Explanation's effect on Trust and Fairness

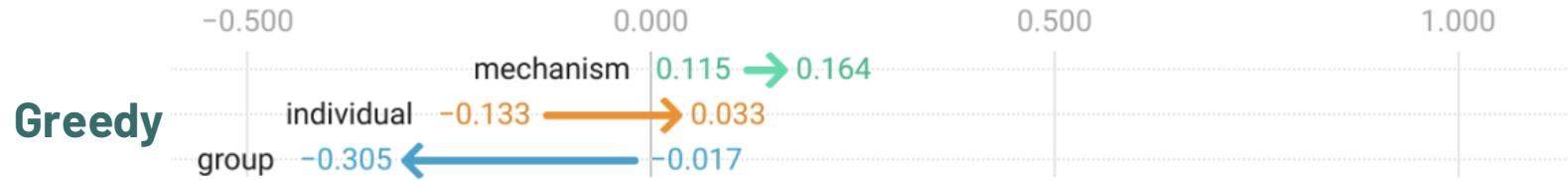
Perceived Trustworthiness



To increase trustworthiness of an algorithm, **explaining the mechanism (inner-working)** is the most effective way



Perceived Fairness



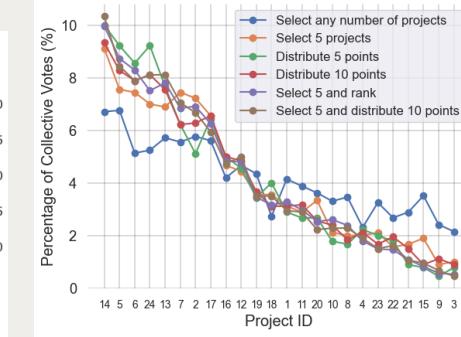
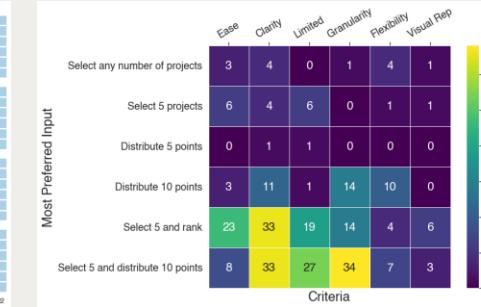
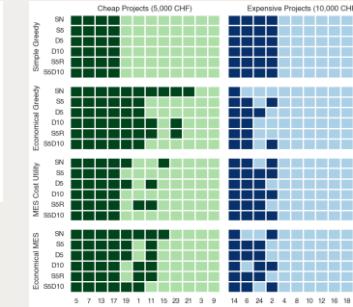
But the increased trustworthiness **does not** always translate into fairness



Design choices for Stadtidee Aarau 2023

Insights from Lab experiment

ID	Project	SN	S5	D5	D10	SSR	SSD10
17	Bike Lanes on Seefeldstrasse 🚲	✓	✓	2 pts	4 pts	#1	4 pts
16	Multicultural Festival at Sechseläutenplatz 🎉	✓	✓	1 pts	3 pts	#2	3 pts
6	More Night Buses to Oerlikon 🚍	✓	✓	1 pts	1 pts	#3	1 pts
7	Free Open Bad Space in Wollishofen 🌴	✓	✓	1 pts	1 pts	#4	1 pts
14	More Trees in Bellevue & Sechseläutenplatz 🌳	✓	✓		1 pts	#5	1 pts
22	Sustainable Cooking Workshop with Kids 🍽️	✓					
12	Car Sharing System for Young People 🚗	✓					
5	Safe Bike Paths around Irchel Park 🚲	✓					



10-point distribution

+

Method of Equal Shares

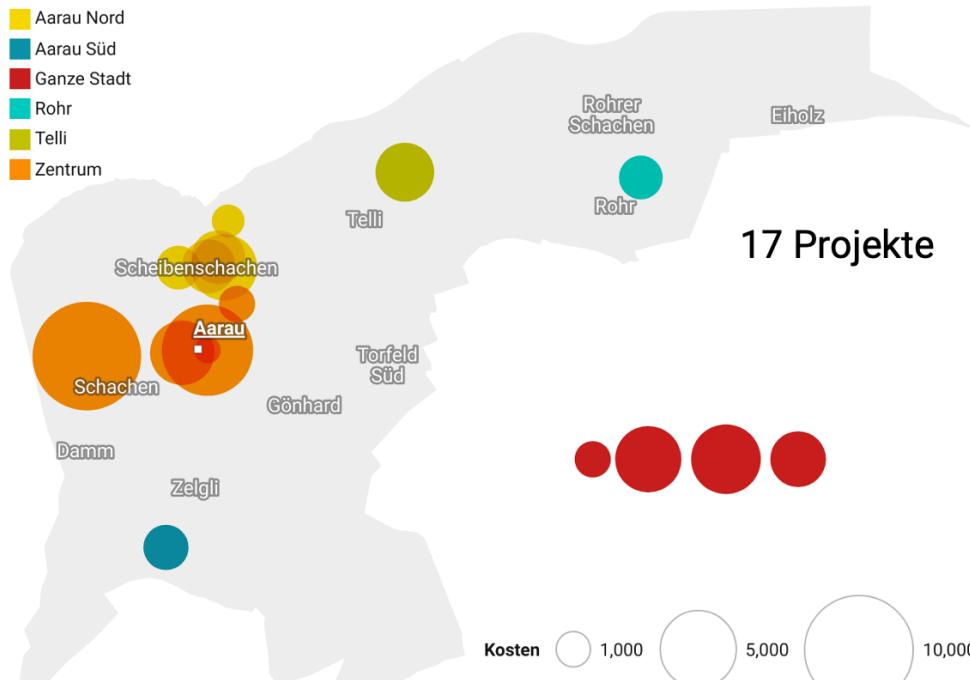
+

Mechanism-oriented
pre-explaination

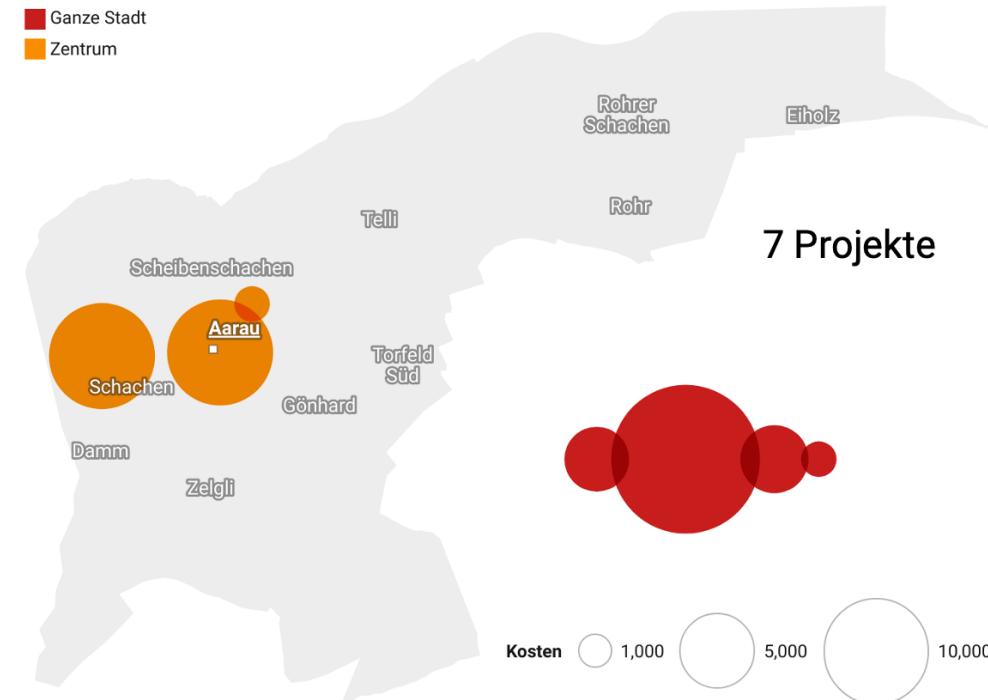
Outcome-oriented
post-explaination

Practical insight: Stadtidee Aarau 2023

Outcome using our method:



If they had used a traditional approach:

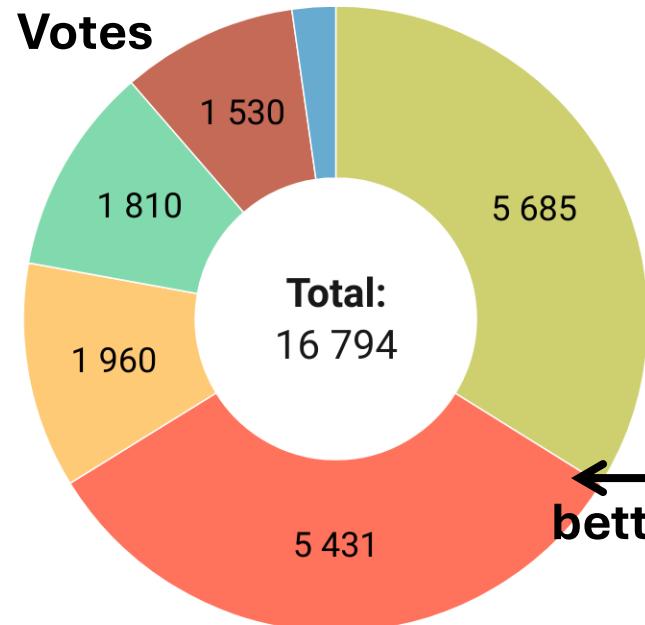


Im Vergleich mit der Standardmethode ist das Ergebnis mit der Methode der gleichen Anteile weniger zentralisiert. Das Budget wird breiter verteilt. Auch kleinere Projekte in der Nachbarschaft haben eine Chance, das Budget zu gewinnen.

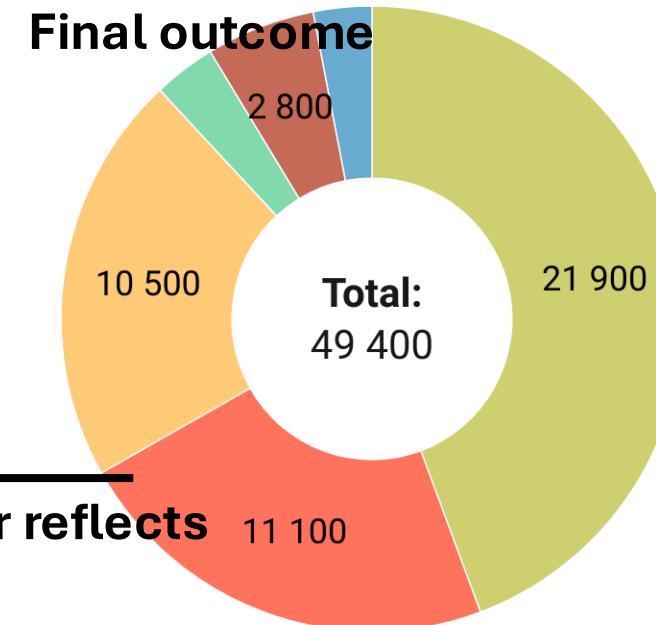
Quelle: Aarau Stadtidee 2023 mit Forschungsprojektes Nationalen Forschungsprogramms 77 (ETHZ, UNIFR, ULeeds).
Diagramm erstellt von Joshua C. Yang (joyang@ethz.ch). Mehr zur Methode der gleichen Anteile unter <https://equalshares.net/>

Practical insight: Stadtidee Aarau 2023

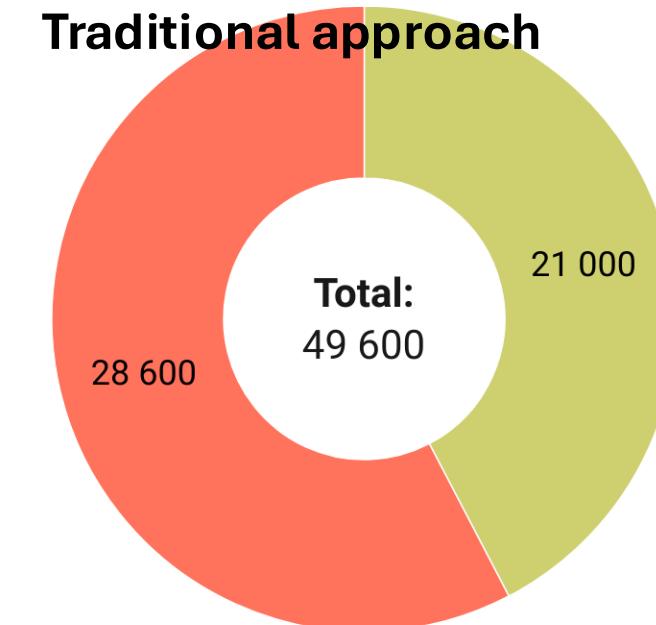
Zentrum Ganze Stadt Aarau Nord Aarau Süd Telli Rohr



Gesamtpunkte für Projekte im
Stadtteil



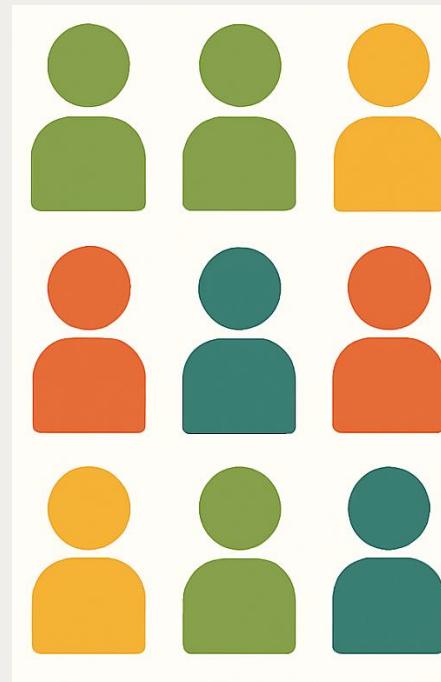
Budget für Projekte im Stadtteil
(Stadtidee mit gleichen Anteilen)



Budget für Projekte im Stadtteil
(Standardmethode)

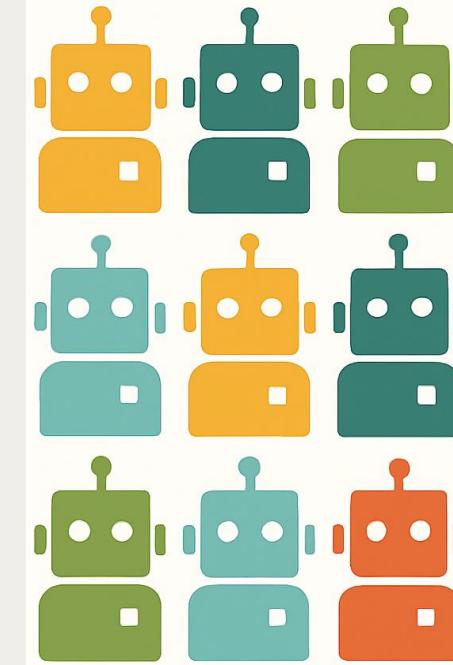
Grafik: Forschungsprojektes Nationalen Forschungsprogramms 77 (ETHZ, UNIFR, ULeeds). Diagramm erstellt von Joshua C. Yang (joyang@ethz.ch).
Mehr zur Methode der gleichen Anteile unter <https://equalshares.net/> • Erstellt mit Datawrapper

In this era of increasing AI involvement in decision-making, the natural question is: **what might be the implication of the use of AI in collective decision-making and democracy?**

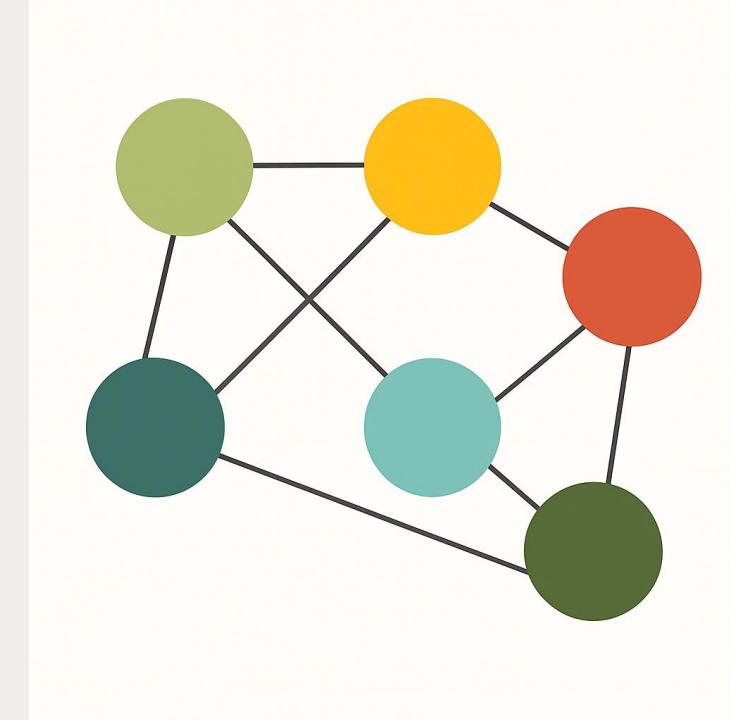


That's too
much voting!

Maybe we can
help?



LLM Voting



Yang, J. C., Korecki, M., Dailisan, D., Hausladen, C. I., & Helbing, D. (2024). ***LLM Voting: Human Choices and AI Collective Decision Making.*** 7th AAAI Conference on AI, Ethics, and Society.

Background: Autonomous Agents and Simulations in LLM



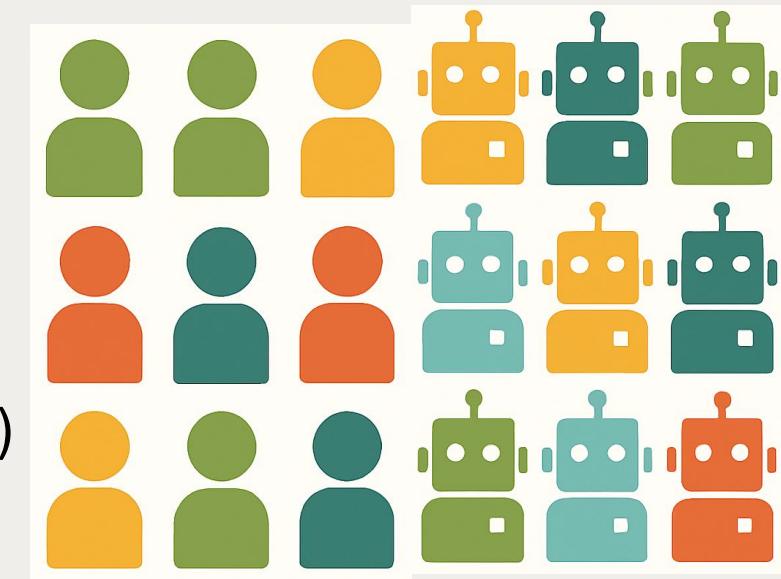
Figure 1: Generative agents create believable simulacra of human behavior for interactive applications. In this work, we demonstrate generative agents by populating a sandbox environment, reminiscent of The Sims, with twenty-five agents. Users can observe and intervene as agents they plan their days, share news, form relationships, and coordinate group activities.

Joon Sung Park, Joseph O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. 2023.

Generative Agents: Interactive Simulacra of Human Behavior. UIST '23. ACM
<https://doi.org/10.1145/3586183.3606763>

Experimental set-up

- **Models tested:** GPT-4 Turbo and LLaMA-2 70B
- **Setup:** 180 LLM agents mirroring the human experiment sample size
- **Task:** Voting on 24 urban projects in Zurich with CHF 60,000 budget
- **Voting methods:** Approval, 5-Approval, Cumulative (10 points), and Ranked voting
- **Interventions tested:**
 - **Temperature** variations (0-2.0)
 - **Persona additions** (based on human survey data)
 - **Chain-of-Thought** (CoT) prompting
 - **List presentation** order effects



Comparing Human and AI votes

- **Aggregated Preferences**

Kendall's Tau: Compares ranking order between LLM and human group outcomes.

$$\tau = Kendall(Rank_{LLM}, Rank_{human})$$

- **Individual Vote Comparison**

Jaccard Similarity Index: Measures overlap in selected projects A between individual i LLM and human voters.

$$J_i = \frac{|A_i^{LLM} \cap A_i^{human}|}{|A_i^{LLM} \cup A_i^{human}|}$$

- **Preference Diversity**

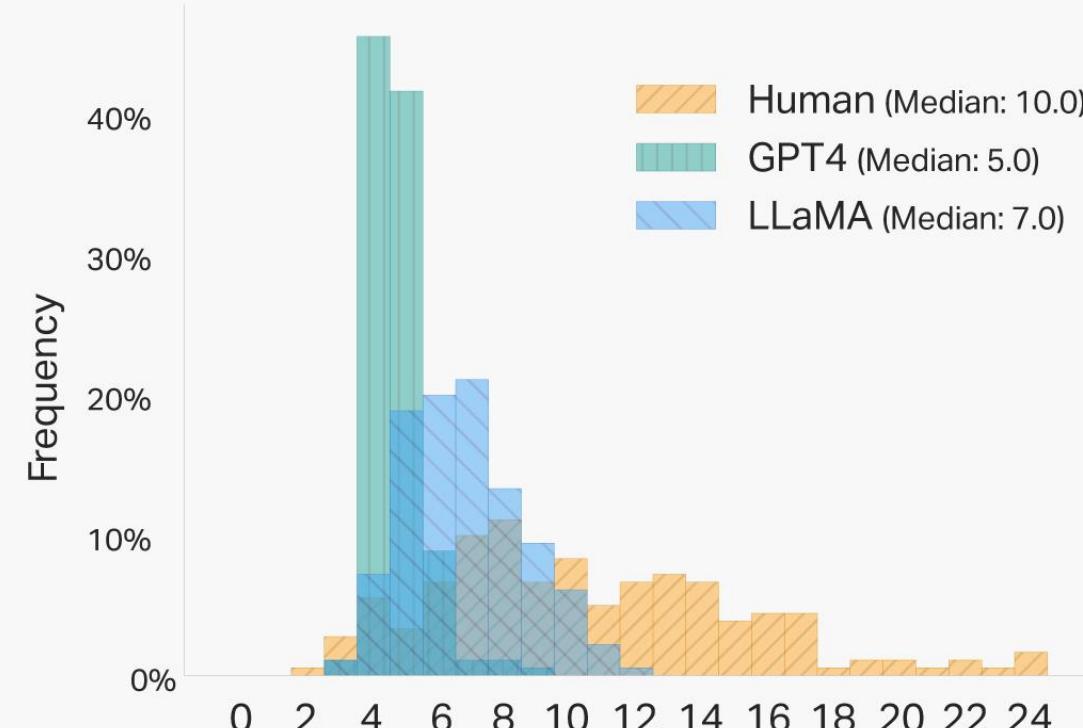
Average Jaccard Distance: Captures diversity of preferences within each group.

$$D_J^g = \frac{1}{\binom{n_g}{2}} \sum_{i < j} 1 - \frac{|A_i^g \cap A_j^g|}{|A_i^g \cup A_j^g|}, \text{ where } g \in \{\text{LLM, human}\}$$

Result

Humans choose with greater variety than AI language models

Number of projects selected in approval voting

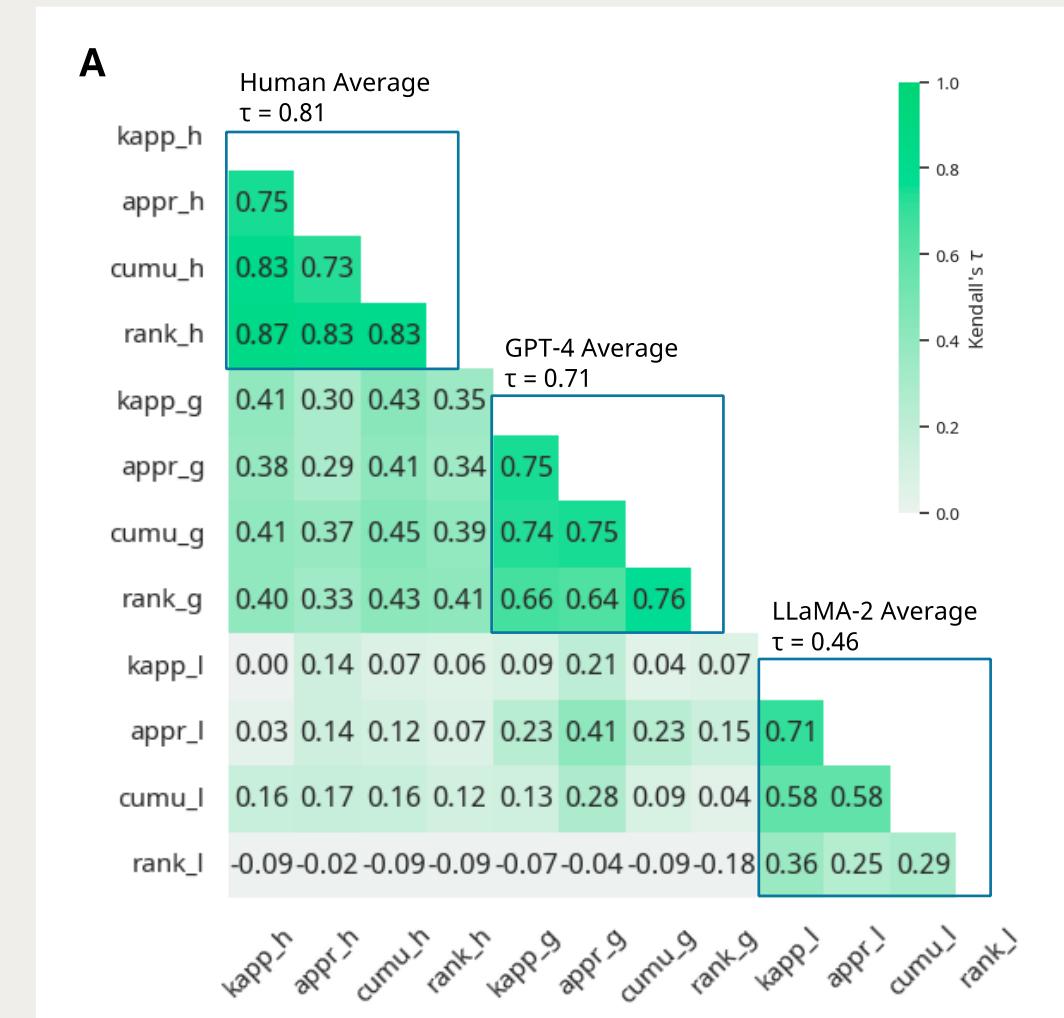


SWI swissinfo.ch

Source: ETH Zurich

Result: The Choice of Voting Methods affects LLMs

- **Humans:** High consistency ($\tau = 0.81$) across different voting methods
- **GPT-4:** Good consistency ($\tau = 0.71$), best in cumulative voting
- **LLaMA-2:** Lower consistency ($\tau = 0.45$), worst in ranked voting



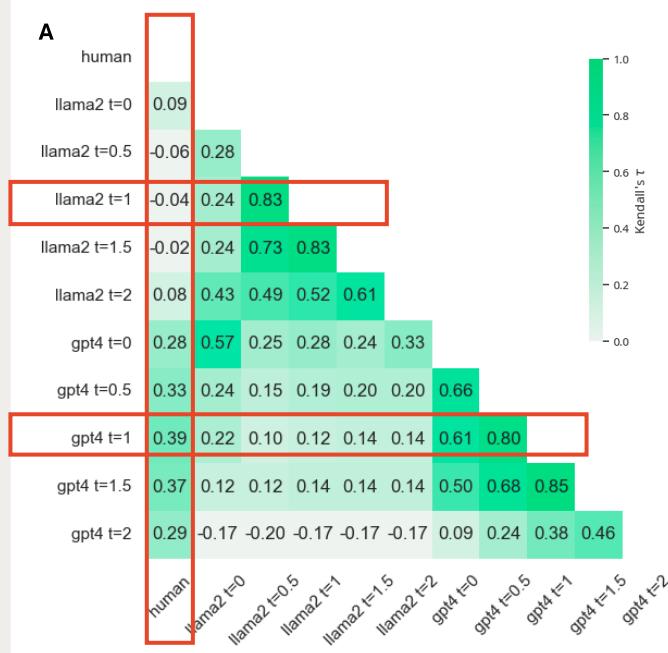
Result: List Presentation Influences LLMs Significantly

B

Rank	LLaMA-2			GPT-4		
	Baseline ($\tau = 1.00$)	Rev Order ($\tau = -0.20$)	Rev IDs ($\tau = 0.19$)	Baseline ($\tau = 1.00$)	Rev Order ($\tau = 0.07$)	Rev IDs ($\tau = -0.25$)
1	#17	#24	#23	1	#5	#17
2	#3	#8	#17	2	#17	#5
3	#13	#10	#3	3	#11	#24
4	#5	#4	#10	4	#24	#1
5	#24	#12	#19	5	#6	#7
6	#23	#18	#1	6	#23	#8
7	#9	#2	#5	7	#7	#14
8	#7	#1	#8	8	#16	#15
9	#21	#14	#15	9	#12	#9
10	#6	#5	#18	10	#10	#11
11	#22	#9	#7	11	#2	#20
12	#15	#3	#13	12	#18	#21
13	#1	#20	#4	13	#4	#23
14	#14	#6	#6	14	#8	#10
15	#10	#16	#12	15	#14	#3
16	#16	#17	#14	16	#22	#13
17	#12	#23	#16	17	#19	#22
18	#11	#13	#24	18	#21	#19
19	#19	#7	#2	19	#1	#2
20	#20	#22	#9	20	#3	#18
21	#8	#19	#21	21	#9	#6
22	#2	#11	#20	22	#13	#12
23	#4	#15	#22	23	#15	#16
24	#18	#21	#11	24	#20	#4

- **LLMs:** Collective outcomes are significantly impacted by changes in order and ID sequence.
- **Humans:** No significant correlation (p -value = 0.66) between order and votes for human voters

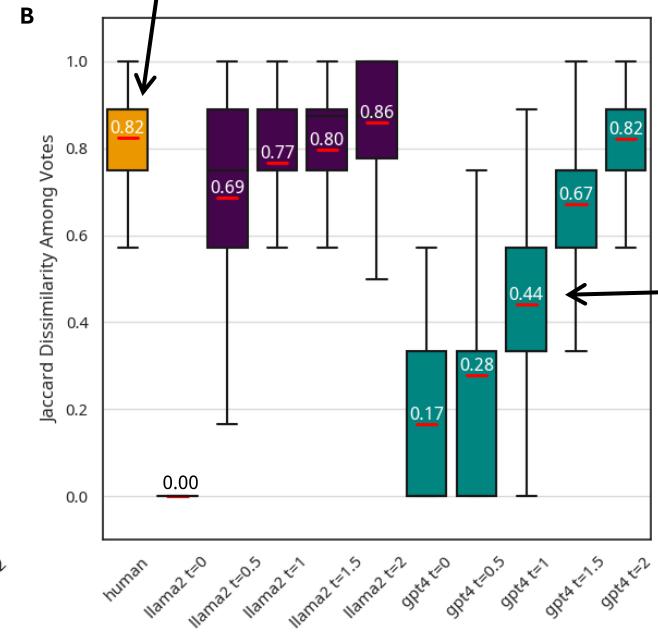
Result: Trade-off Between Alignment and Diversity



Best Alignment at temp = 1

LLM voting patterns are most aligned with human voters.

Human choices are very diverse



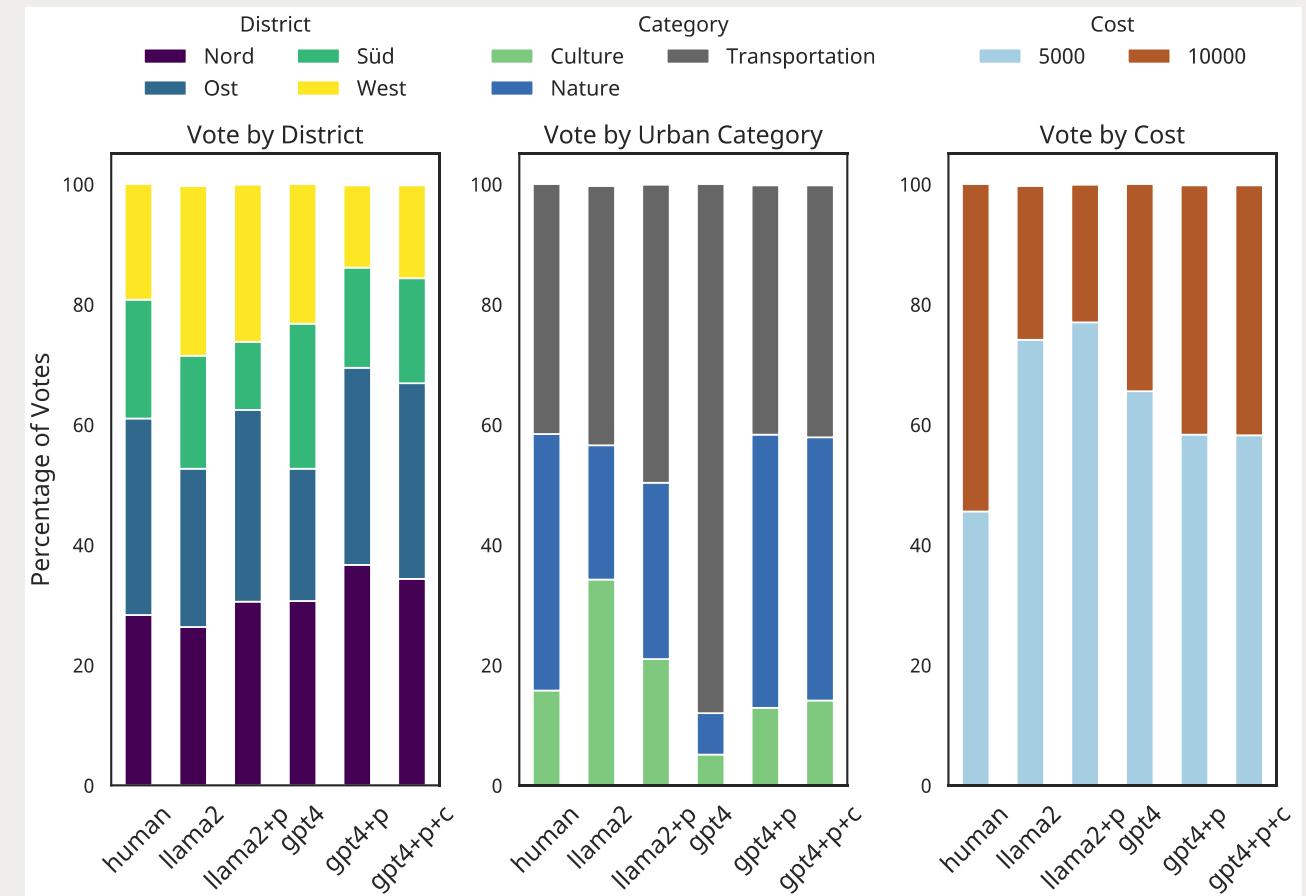
GPT4 diversity depends highly on the temperature

Human Diversity at temp = 2

LLMs match human level of diversity.

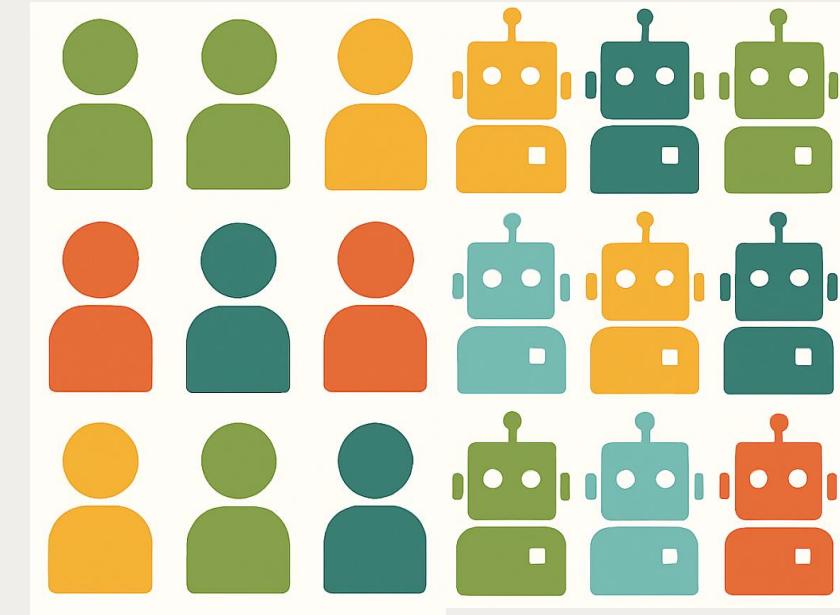
Result: Persona Addition

- Improved individual vote alignment: LLaMA-2 ($0.14 \rightarrow 0.21$), GPT-4 ($0.18 \rightarrow 0.30$)
- GPT-4 with persona achieved highest collective alignment ($\tau = 0.54$)
- **LLMs became overly focused on self-interest compared to humans**



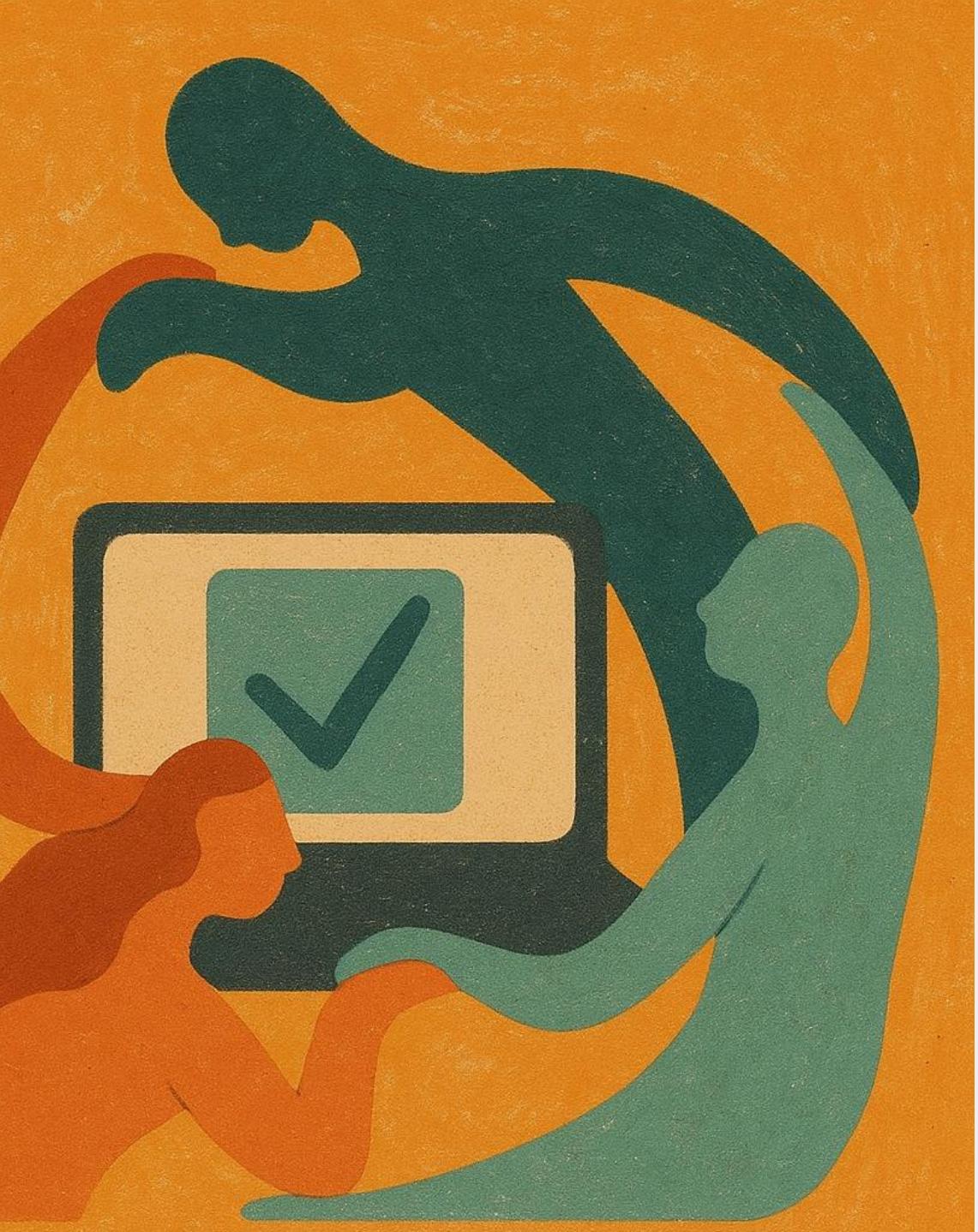
Takeaway

- AI voting was uniform, while human preferences were diverse.
- AI was influenced by irrelevant factors like option order.
- GPT4 quantified preference intensity well, but that's a technical strength, not a democratic one.
- **Conclusion: Support, Not Replace:** AI may assist with summarising, organising, and informing, but it lacks the civic engagement, accountability, and deliberative transformation that democracy fundamentally requires.



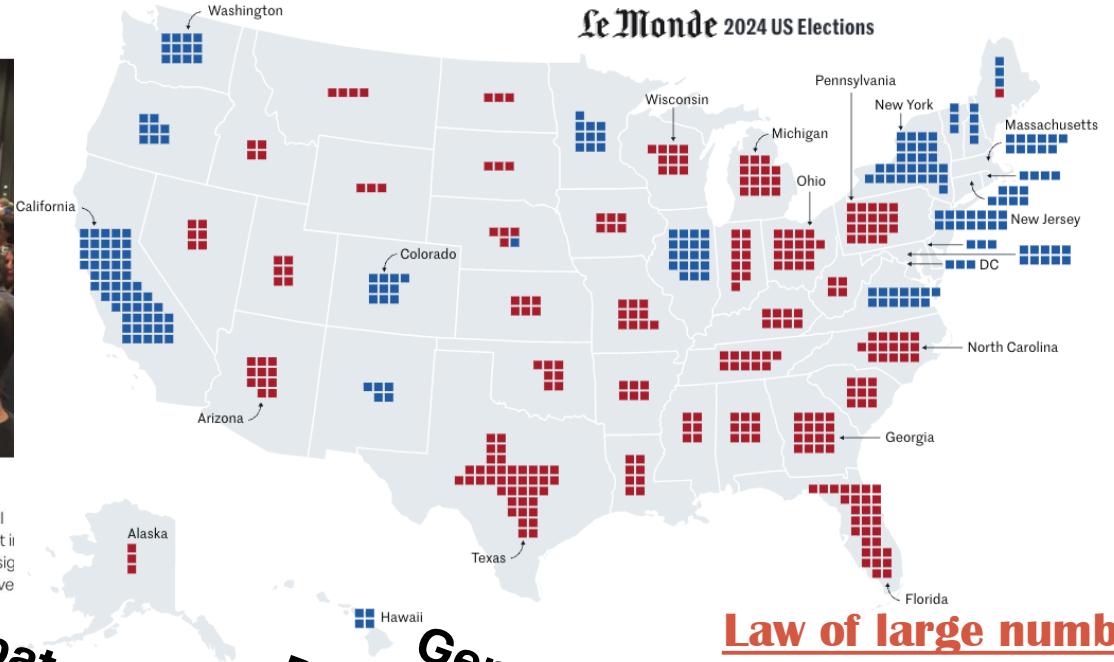
Perhaps a productive question is: **how can we design algorithmic processes that help humans become better democratic participants?**



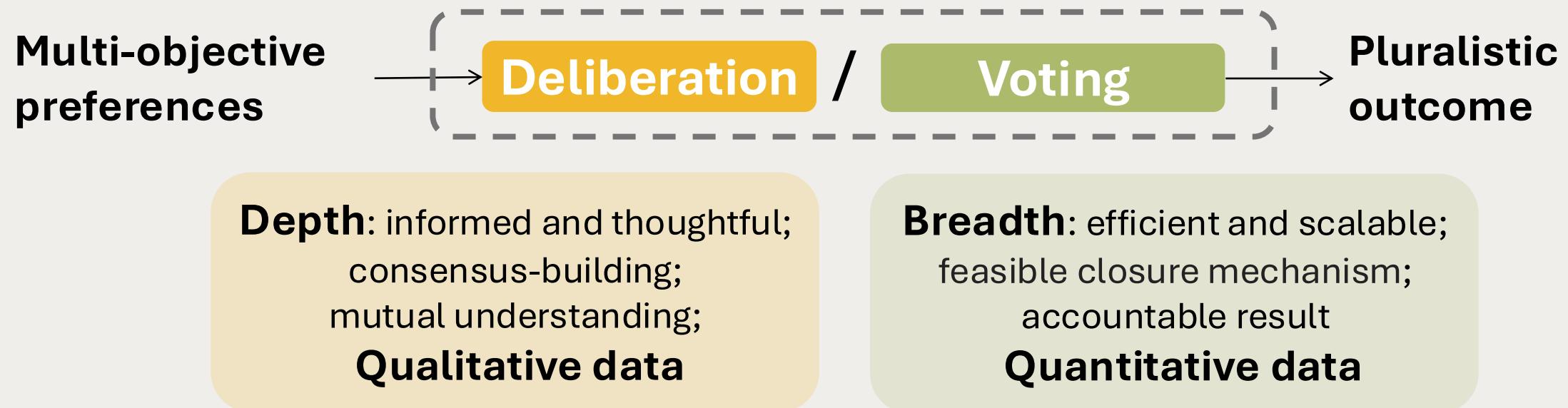


Bridging Voting and Deliberation

Tension between deliberation & voting



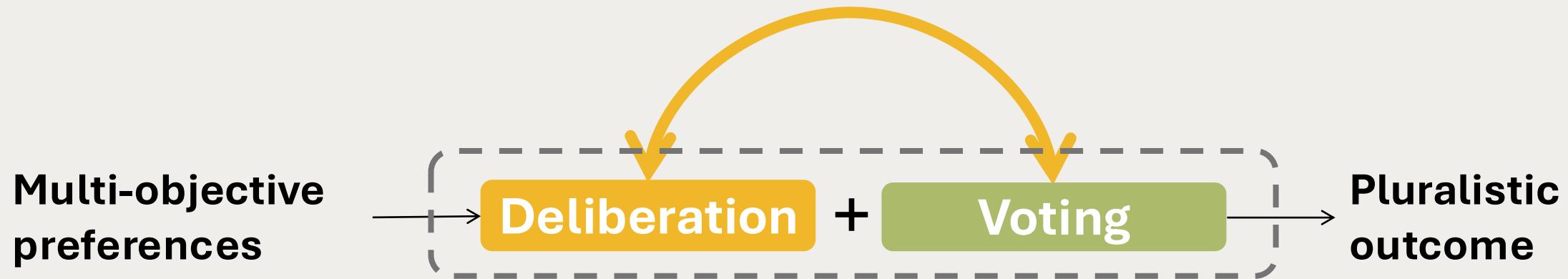
Trade-off between deliberation & voting



Chambers, S., Warren, M.E. Why Deliberation and Voting Belong Together. *Res Publica* (2023).

Practical Questions to be Answered

How can algorithmic methods support hybrid democratic processes?



How can voting guide deliberations?

Proposed digitally supported framework



- Preserves the **independence** and diversity conditions by allowing people to develop their own views first
- **Quantitative data** for algorithmic methods to support deliberation

Methodology: Participatory Action Research

- Collaborative **co-design** with organisers and participants
- Embedded, iterative process development
- **Secondary data analysis** of anonymised datasets collected by organisers
- **Empirical insight** grounded in **real-world constraints and needs**
- The **contribution** lies in the methods co-developed



The 2024 iteration of Kultur Komitee Winterthur

Case Study 1

Kultur Komitee

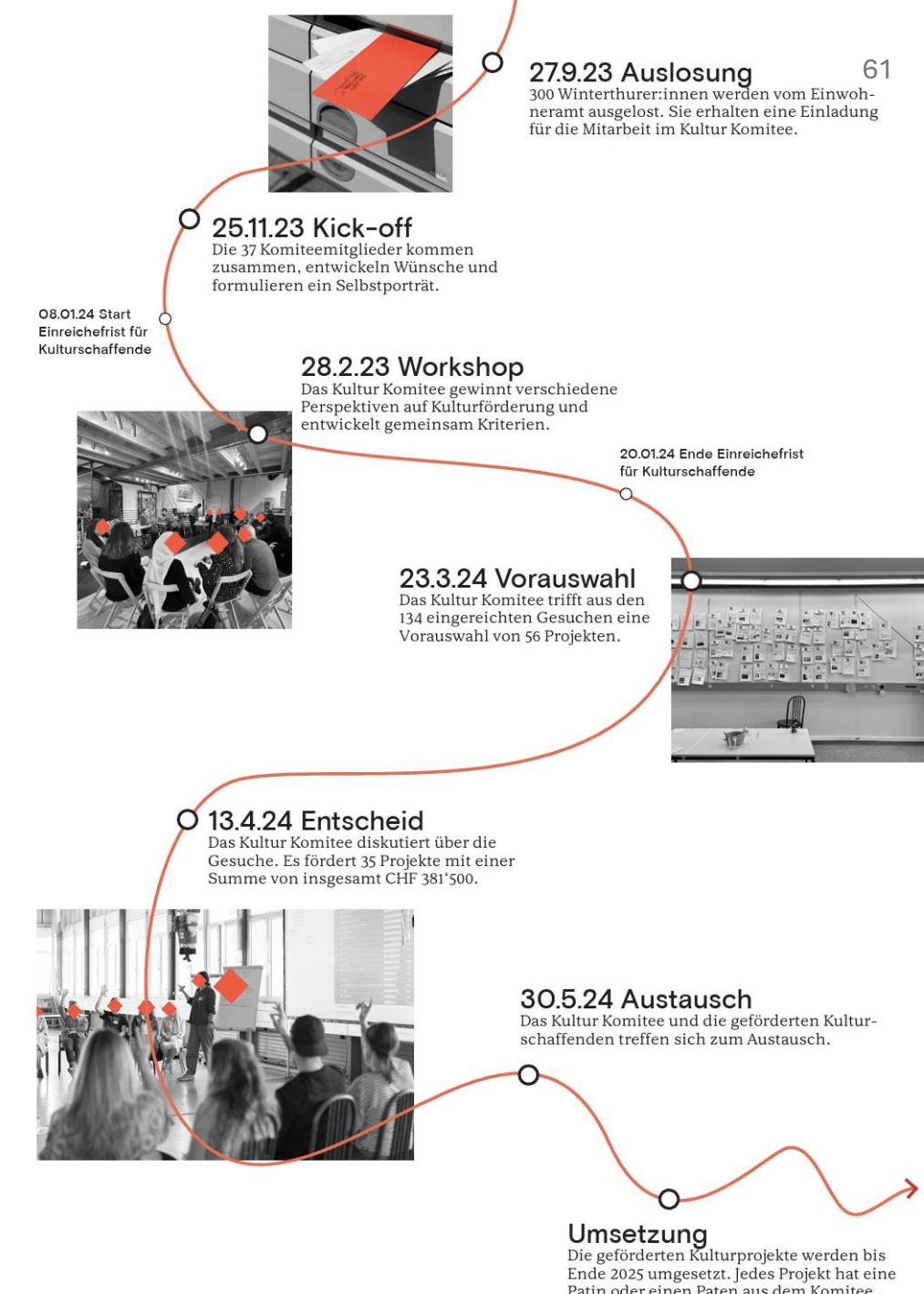
Winterthur



Case: Kultur Komitee Winterthur

- **"Budgeting Assembly"**: Combine Citizen Assembly with Participatory Budgeting
- 400,000 CHF every year to democratise art and culture
- **Sortition**: Around 40 randomly invited citizens come together need to decide how this budget should be spent
- Between 2021 and 2024, a total of 107 cultural projects were funded with a total of CHF 1.2 million.

Stiftung für Kunst,
Kultur und Geschichte



Step 0. Invite Citizens to Form Committee

Using the idea of sortition, the organising team sends out 200 letters, inviting random citizens to form a “Culture Committee”. Around 30-40 citizens would end up joining in this process, participating online and in workshops.



Step 1. Committee Identifies Core Criteria

Committee come up with normative goals (a topic, a theme) and use these as criteria in the call for projects

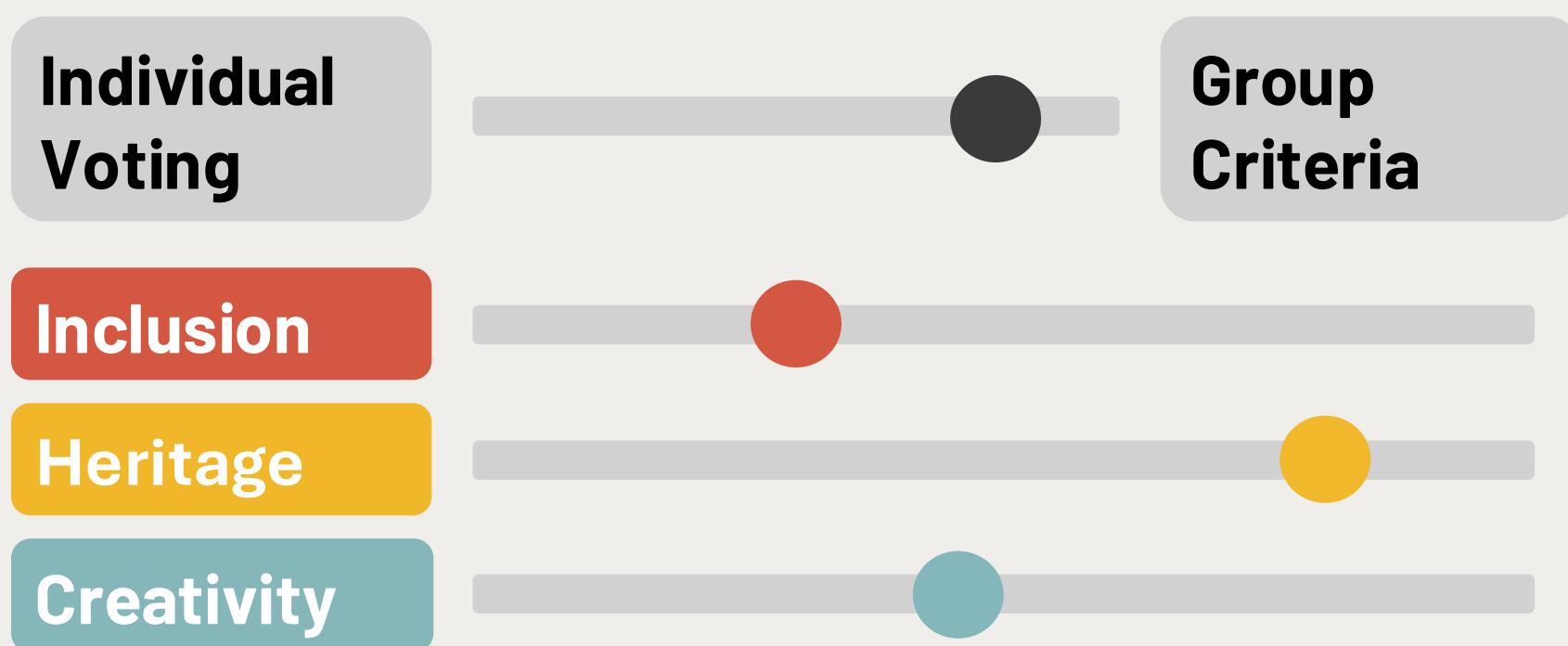


Step 2. Committee Co-tuning Algorithm

Using an online vote, the committee decides

- (1) The **split** between the group deliberation and voting,
- (2) The **relative importance** of each value

These values would become the core parameters of the Komitee Equal Shares algorithms



Gewichtung der Stimmen

1

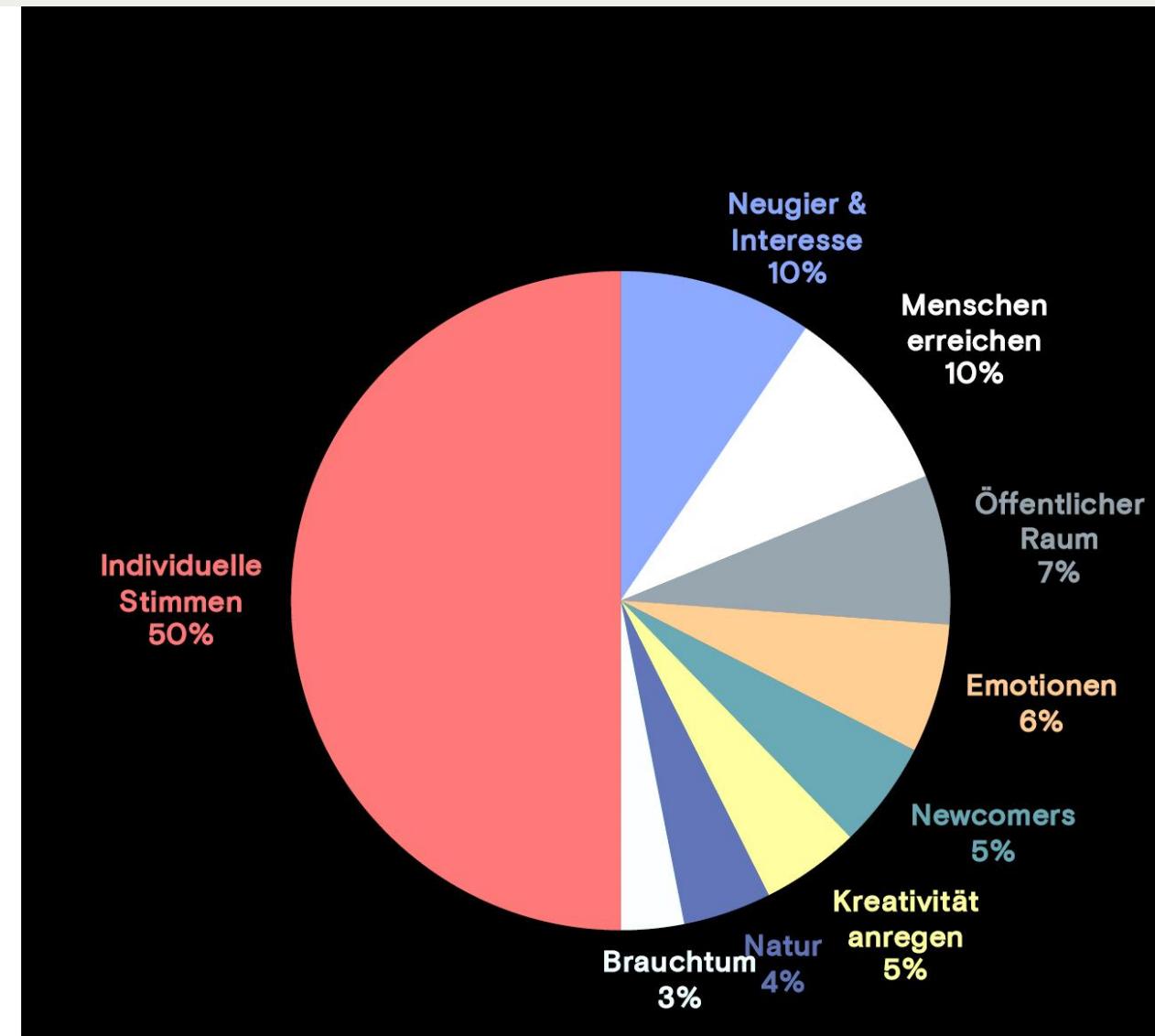
Wirkungsfelder

	Prozent	Anzahl Stimmen	Rang
Neugier & Interesse	19.0%	114	1
Menschen erreichen	18.7%	112	2
Öffentlicher Raum	14.7%	88	3
Emotionen	12.7%	76	4
Newcomers	10.5%	63	5
Kreativität anregen	9.7%	58	6
Natur	8.7%	52	7
Brauchtum	6.2%	37	8
Total	100%	600	

2

Individuelle Stimmen

Wirkungsfelder	Individuelle Bewertung
1971	2029
49%	51%



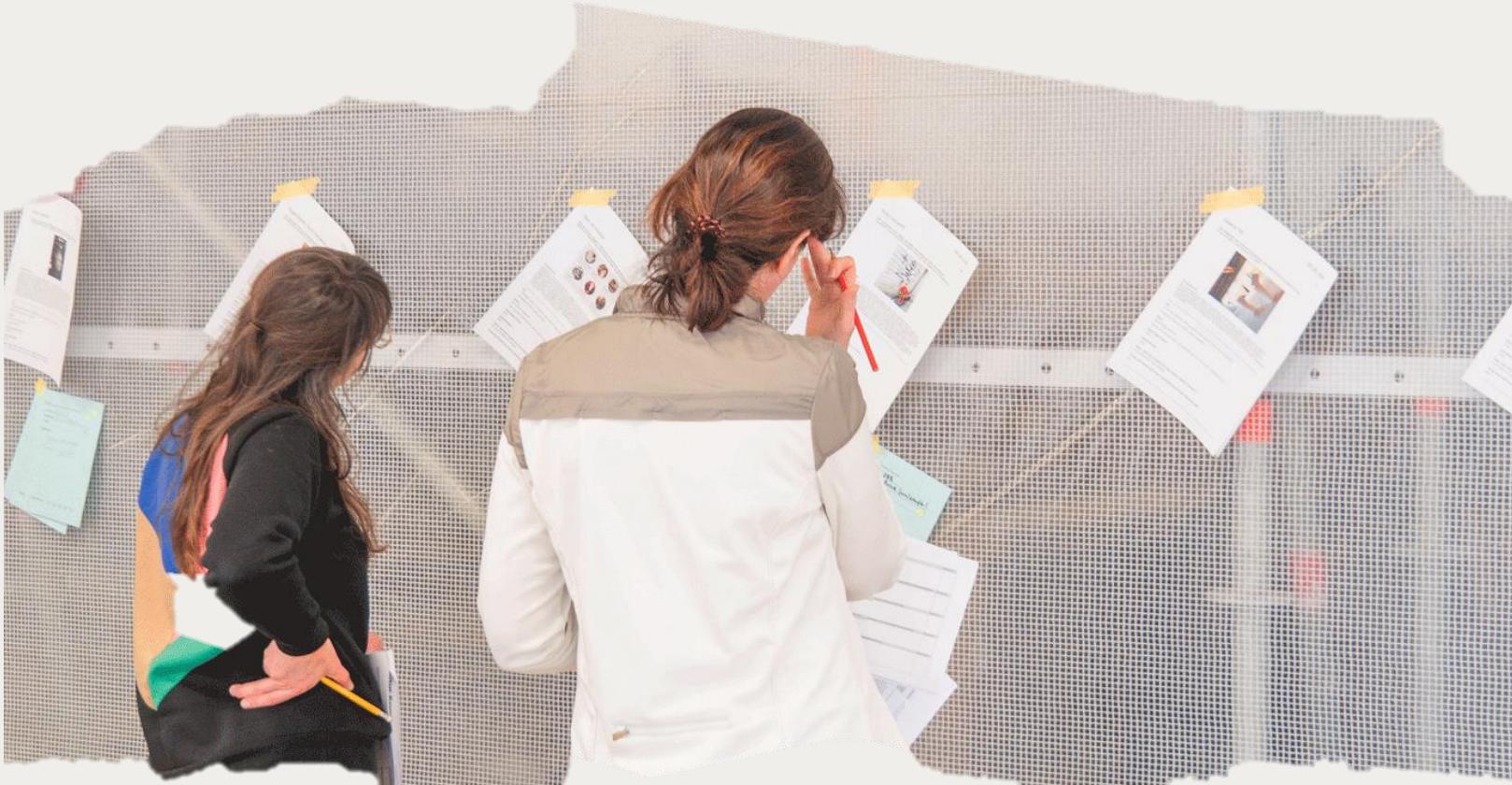
Step 3. Criteria Groups Deliberate

The participants go to a criteria table in each round to deliberate together and judge the projects based on the particular criterion. The groups have to decide how the points should be allocated to projects. The number of available points in each group correlates to the relative importance of the criterion.



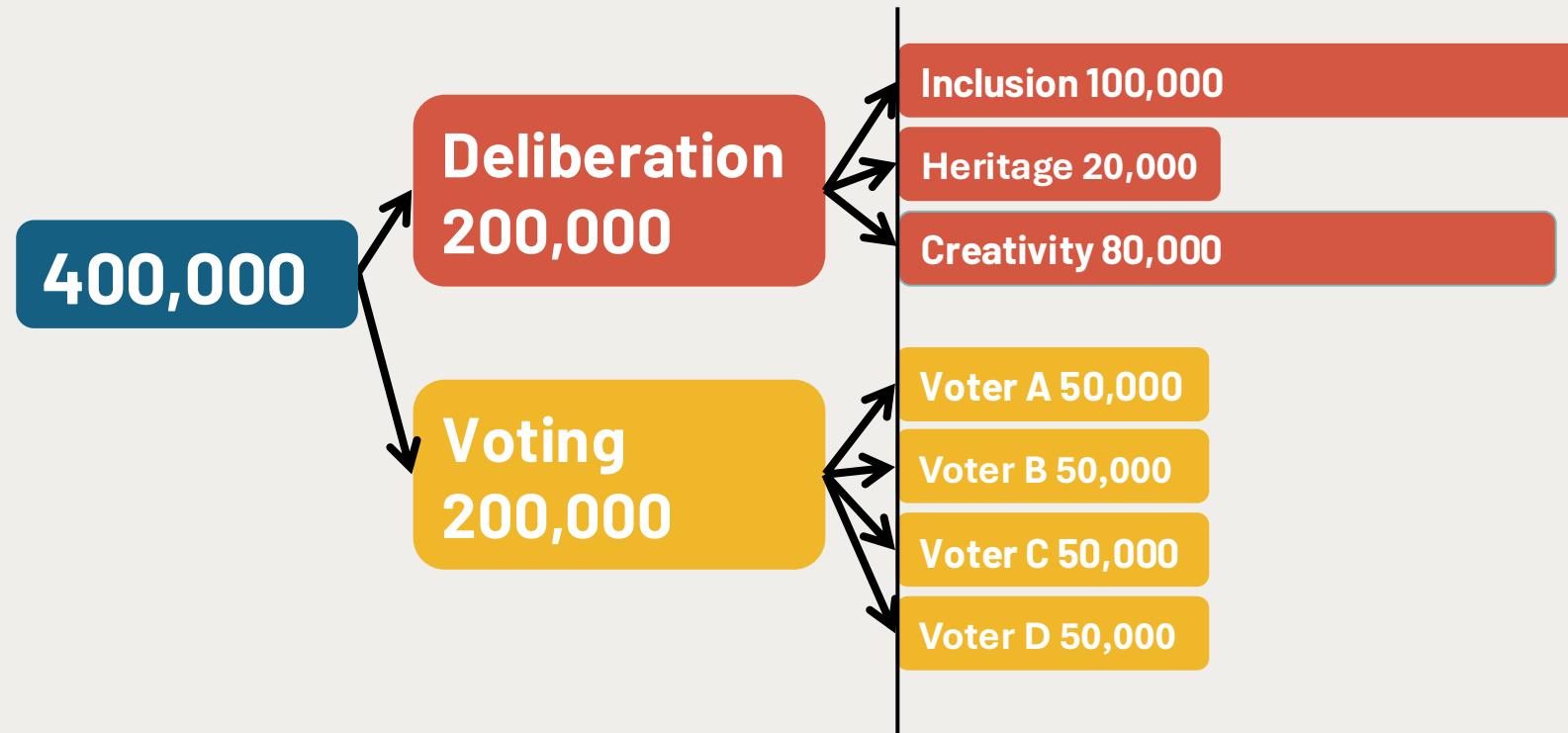
Step 4. Individual Voting

The individual participants go around the room with all the projects present and cast their 20 points to projects that they like.



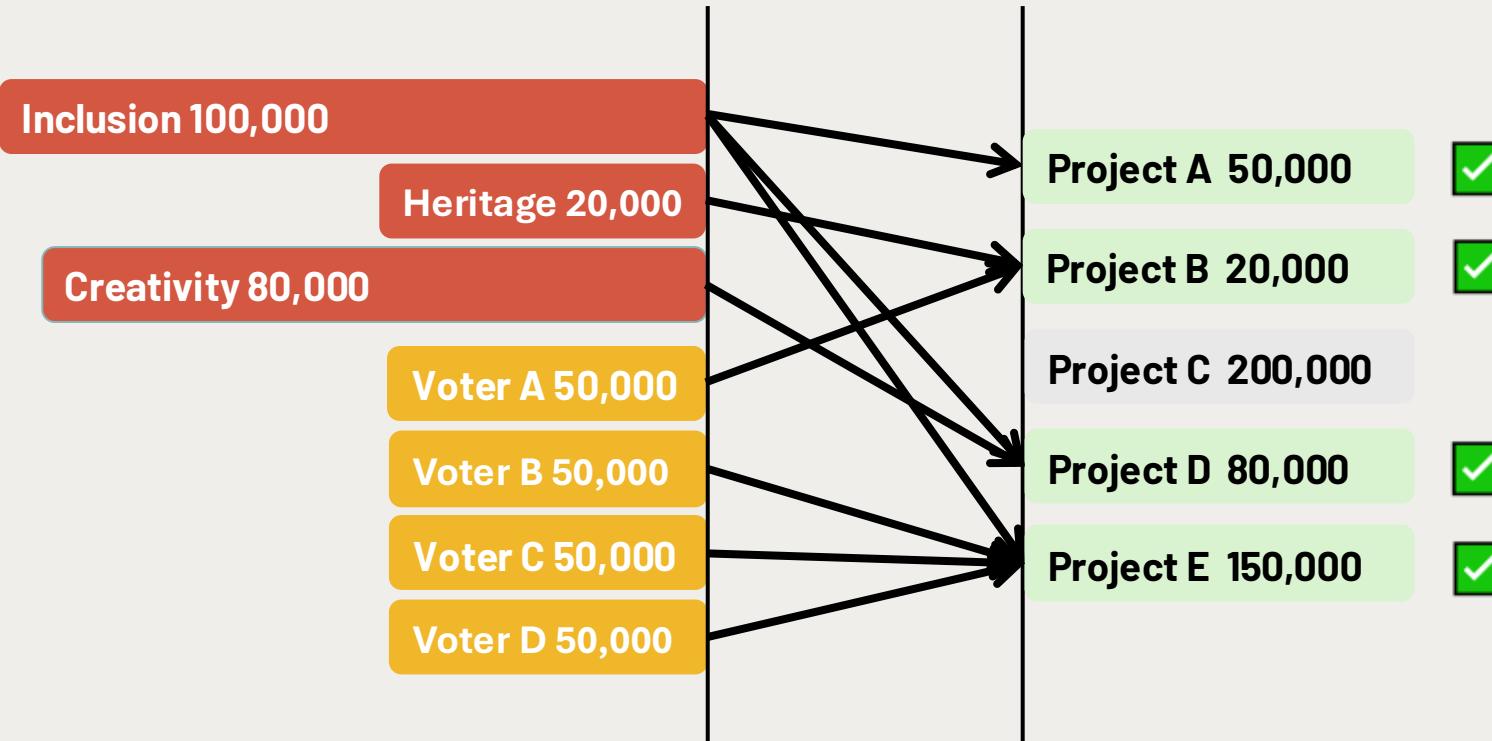
Step 5. Algorithm Splits Virtual Budget

According to parameters decided together by the committee, the Komitee Equal Shares algorithm allocates the virtual budget into many different buckets for calculation.



Step 6. Algorithm Funds Projects

Starting from the most supported, the algorithm (using a MES approach) tries to fund the projects one by one and see if the supporters (groups and individual) have enough collective budget in their buckets to fund them

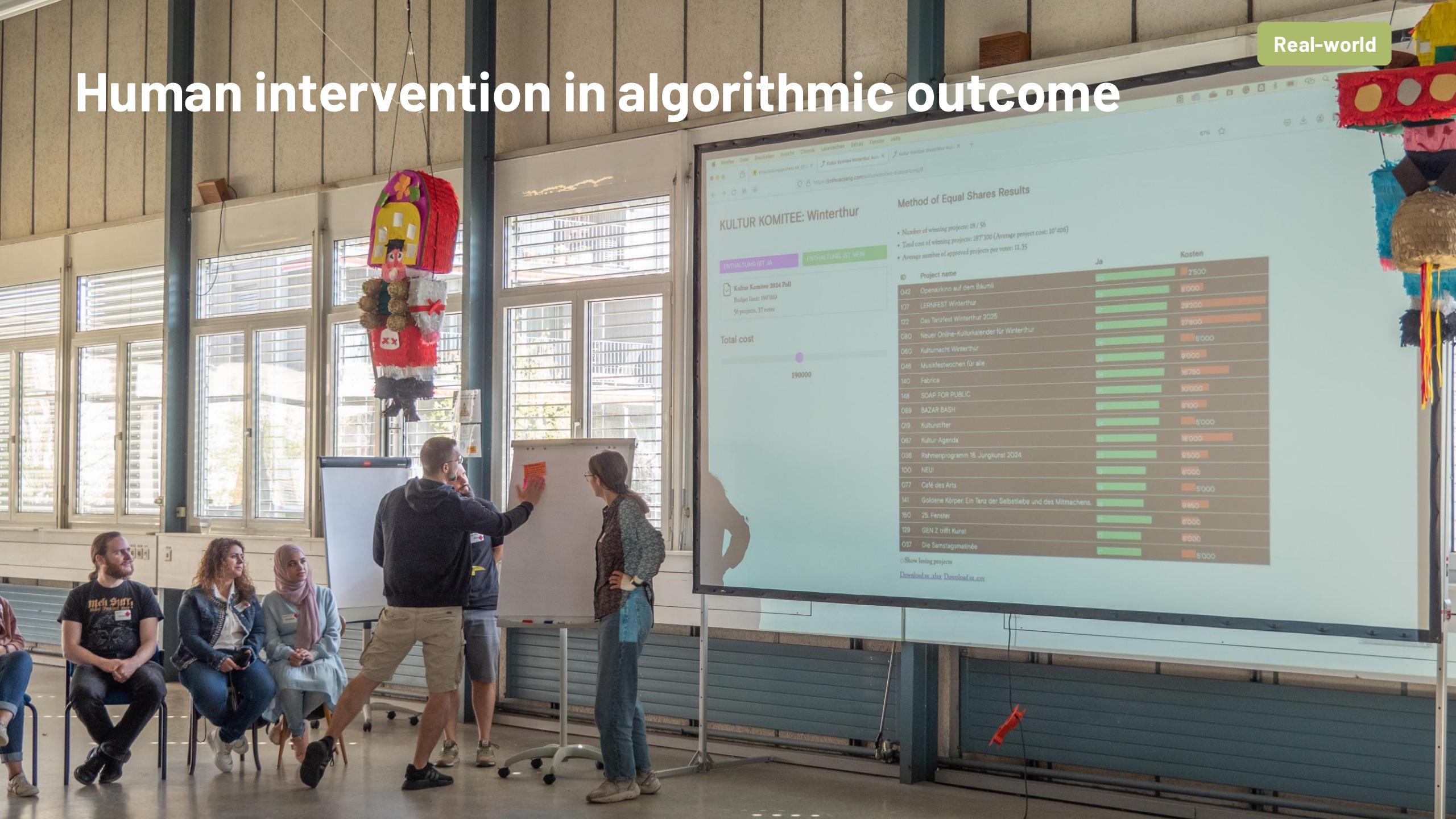


Step 7. Interacting with the Outcome

The interactive outcome table shows exactly how each projects are funded with real votes with group/individual support. Participants could also play around with the slides to adjust total budget and group-individual ratio to produce alternative scenarios.



Human intervention in algorithmic outcome



Method of Equal Shares Results

- Number of winning projects: 18 / 56
- Total cost of winning projects: 187'300 (Average project cost: 10'406)
- Average number of approved projects per voter: 11.35

ID	Project name	Ja	Kosten
042	Openairkino auf dem Bläumli	50	2500
107	LERNFEST Winterthur	28	8'000
122	Das Tanzfest Winterthur 2025	21	28'600
080	Neuer Online-Kulturkalender für Winterthur	21	5'000
060	Kuturnacht Winterthur	20	9'000
048	Musikfestwochen für alle	20	16'750
140	Fabrica	20	10'000
148	SOAP FOR PUBLIC	20	10'000
069	BAZAR BASH	24	9'100
019	Kulturstifter	24	5'000
067	Kultur-Agenda	21	18'000
036	Rahmenprogramm 18. Jungkunst 2024	21	9'500
100	NEU!	19	6'000
077	Café des Arts	19	5'000
141	Goldene Körper: Ein Tanz der Selbstliebe und des Mitmachens.	19	9'650
150	25. Fenster	19	6'000
129	GEN Z trifft Kunst	17	6'000
037	Die Samstagsmatinée	17	5'000

Show losing projects

Download as .xlsx Download as .csv

Step 8. Voting Receipts

When the calculation is done, the interface also produces a receipt that details what the group/voter “spent” on with their votes, essentially how their votes were used.

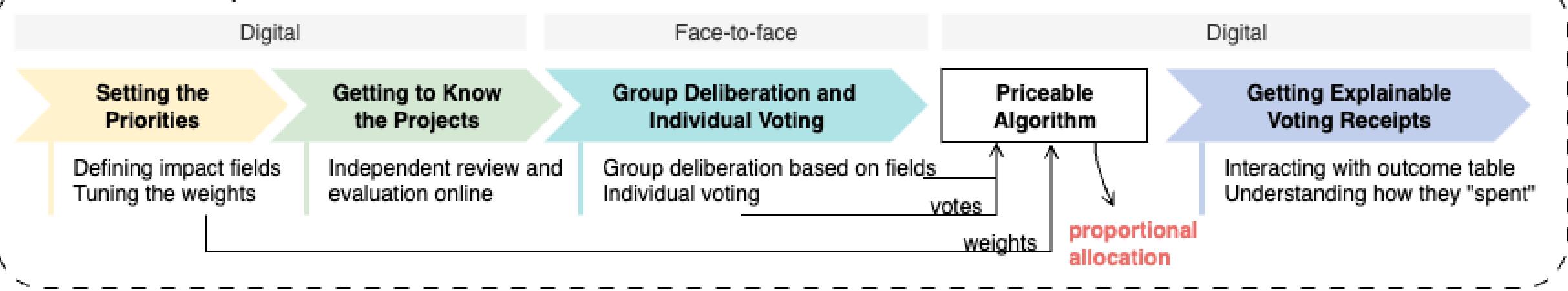
🔍 Neugier		TOTAL
		37,164
ITEM	VOTES	SPENT
Vélo volant – Das Figur...	9	3610
Tier & Text	4	4544
1 Jahr Tüfteln in der Fa...	4	3262
"Silvesterabend" – eine...	3	8435
Winterthurer Druckwoc...	3	1646
Das Winterthurer Publi...	3	0
Die Immigranten / The ...	3	0
Kinder- sowie Jugendt...	2	4684
LES FEMMES FATALES	2	2324
Samstagsmatinée	1	2620
Rauk-it	1	2582
Der Wurm in meinem K...	1	1929
Never Enough	1	1051
Tag der Vielfalt 2025	1	478
Subtotal	38	37,165

❤️ Emotionen		TOTAL
		21,243
ITEM	VOTES	SPENT
Nicht schuld!	5	3177
Eisblumen – ein etwas ...	5	1939
Tuuli	4	5231
Vélo volant – Das Figur...	3	1173
Green Rock	2	2969
Schreibprojekt «Palazz...	2	2866
LES FEMMES FATALES	2	2326
Never Enough	1	1056
Konzert Kammerchor u...	1	506
Ruhe (un)sanft!	1	0
Subtotal	26	21,243

人群		TOTAL
		37,646
ITEM	VOTES	SPENT
Cinema Expanded: A ...	5	8670
Eisblumen – ein etwas ...	4	1560
ANTE Konzertreihe	4	1338
Dialogplatzkonzerte 2...	3	3638
INDUSTRIA FESTIVAL: ...	3	3311
BASS & PIXELS	3	0
Sonntag für alle am Ta...	3	0
The Car Rots (das Aut...	2	5233
KleinKunstRallye Winte...	2	5145
Schreibprojekt «Palazz...	2	2896
Nicht schuld!	2	1289
Vélo volant – Das Figur...	2	788
Schattenwelten (AT)	1	1530
KULTHURPULS – Gesp...	1	1187
Never Enough	1	1062
Subtotal	38	37,647

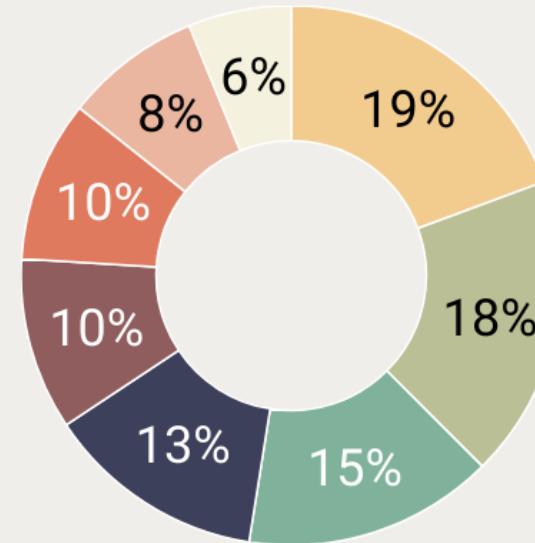
<https://kk25.win/>

The Komitee Equal Share Framework

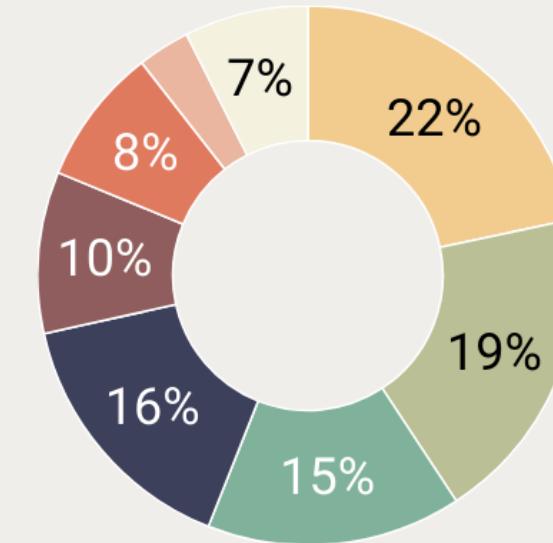


How the final set aligns with the defined values

- █ Curiosity & Learning
- █ Community & Connection
- █ Public Space & Accessibility
- █ Emotional Impact
- █ Support for Newcomers
- █ Audience Creativity
- █ Nature & Environment
- █ Tradition & History



Co-defined
weights



Final KES
Selction

Side observation: Voting emerges in deliberation in the wild



Ranked voting on the floor



Point distribution with sticky notes

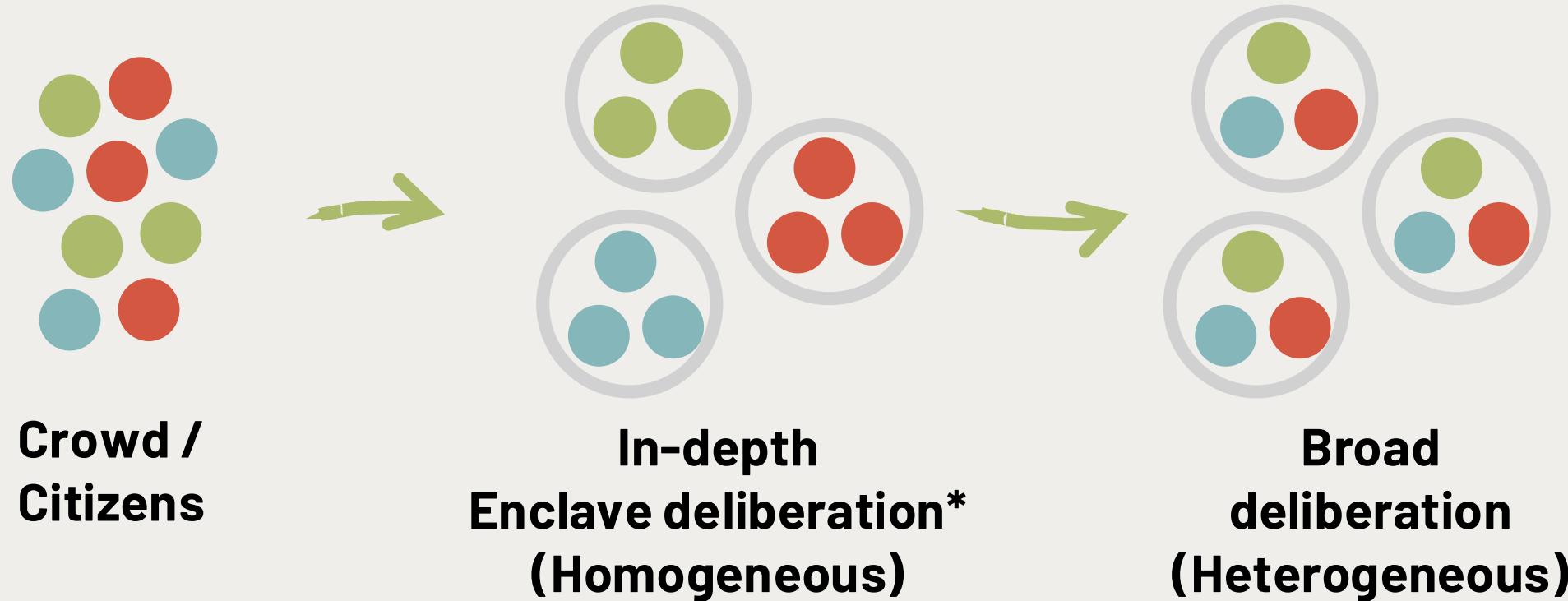
Method 1: Preference-based Clustering for Deliberation (PCD)

Online voting data as input → structured deliberation groups



Context: Enclave deliberation*

Let **like-minded people** talk together first, **before** mixing with others. It helps **marginalised voices** articulate their ideas clearly **without being dominated**.



*Abdullah C. & Karpowitz C. & Raphael C., (2016) "Affinity Groups, Enclave Deliberation, and Equity", *Journal of Public Deliberation* 12(2).

*Fraser, Nancy. (1990). *Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy*.

Preference-based Clustering for Deliberation

From voting patterns → to opinion space → to deliberation groups

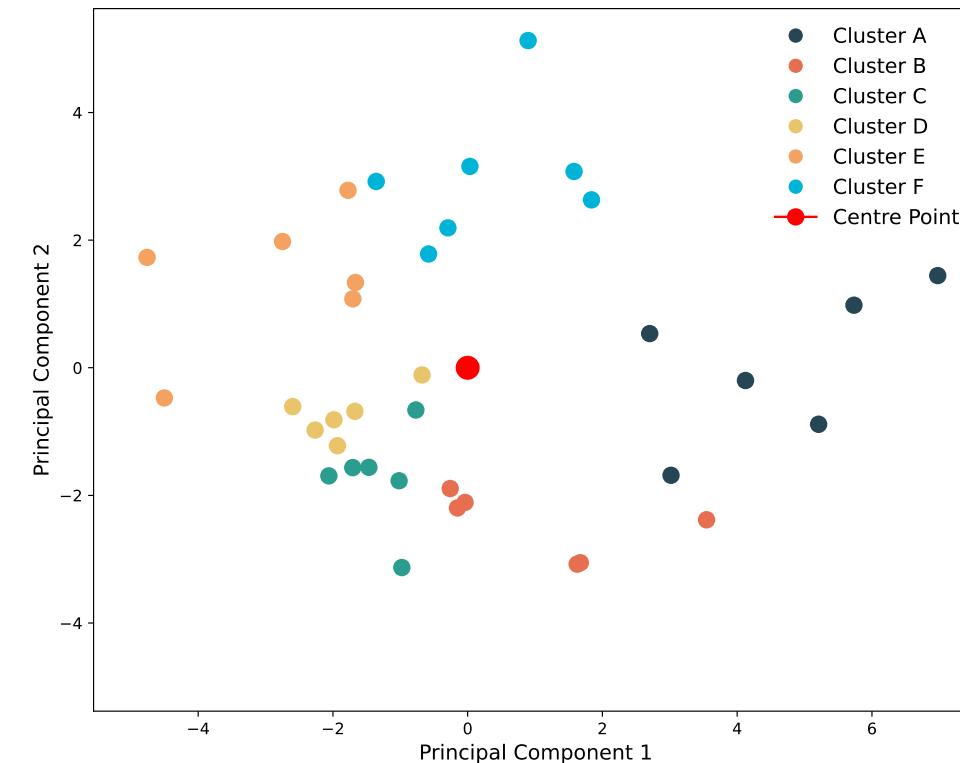
Online vote (direct preferences)

	A	B	C
Voter 1	✓	✗	✓
Voter 2	✗	✓	□
Voter 3	✗	✓	□

Radial Clustering
with Dimension
Reduction



Deliberation groups



Preference-based Clustering for Deliberation

Designed for quick adjustability on the ground



If someone doesn't show up or leaves early, you can **manually redraw the slice boundaries** to keep groups balanced.

=> ideal for **low-tech, real-world deliberation settings** where flexibility is key.

Preference-based Clustering for Deliberation



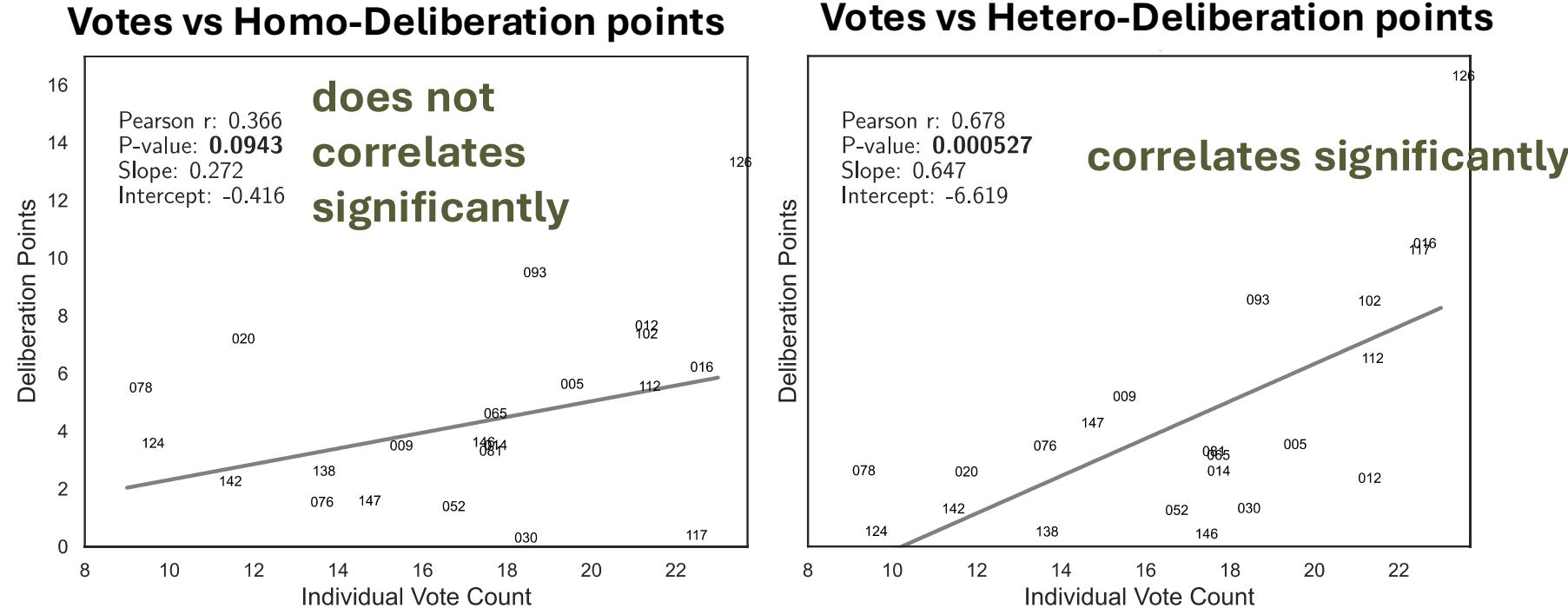
**First round:
homogenous deliberation**



**Second round:
heterogeneous deliberation**

Question: How do deliberation outcomes differ from individual voting patterns?

Result: Hetero Deliberation outcome similar to Voting

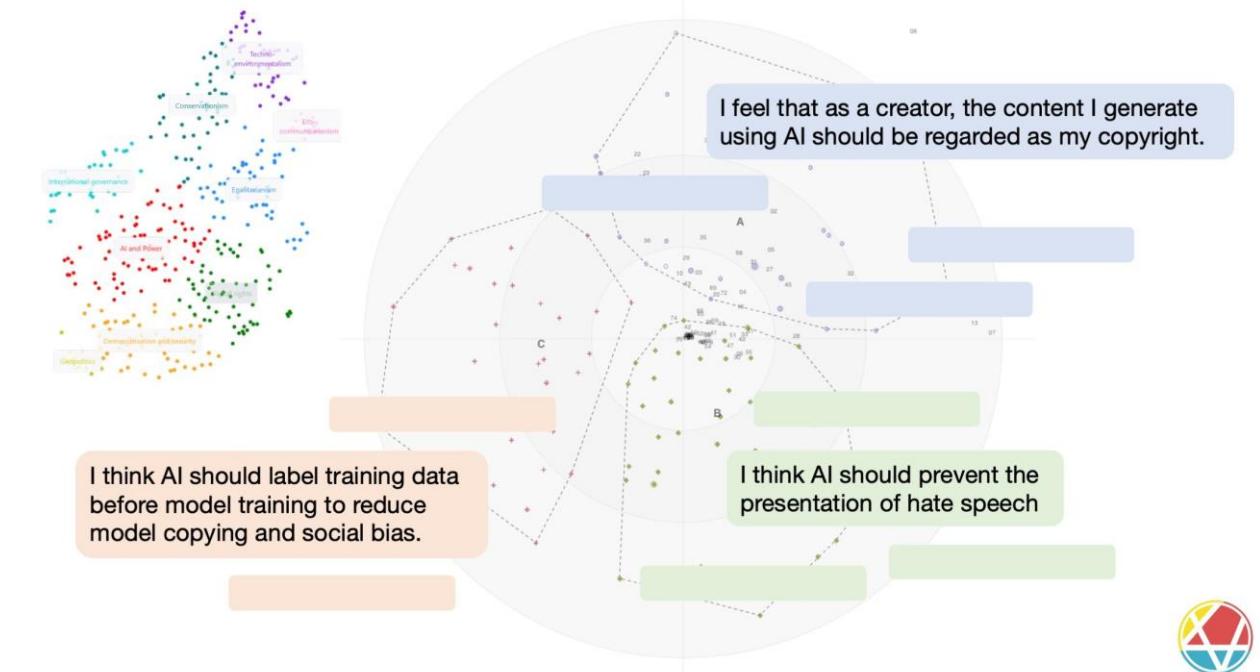


Homogeneous deliberation groups surfaced preferences that were not dominant in voting outcomes.

Some Feedback

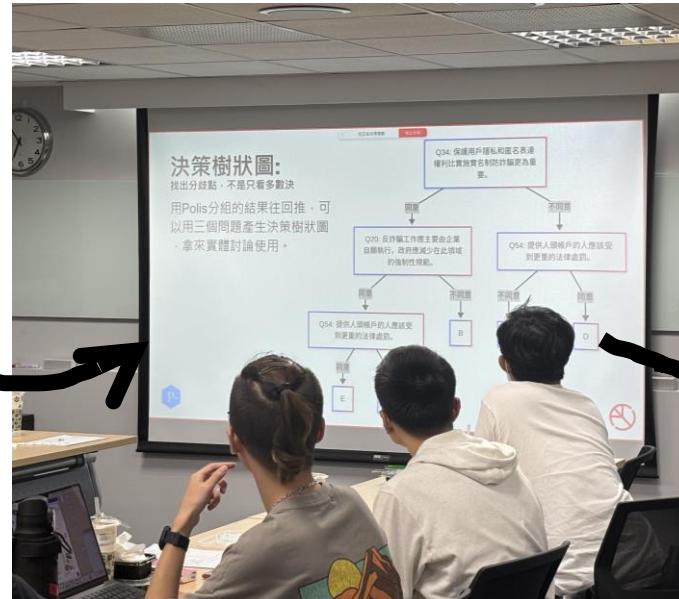
- “It took a lot of pressure off of having to do everything ‘right and fairly.’ That was great.”
- “I think the method is good. People who aren’t familiar with the culture would probably be overwhelmed if they had to decide everything on individual votes.”
- “I got out of my bubble a bit. I’m on the board of the SVP Winterthur, and in my free time, I am in more conservative circles. It was good to work on a project with young, dissenting people again. Especially since I was an active Social Democrat in my youth and was involved in various youth organizations. It was very enriching.”

Case Study 2

 vTaiwan

The vTaiwan Process

STATEMENT	OVERALL 105	A 11	B 43	C 23	D 16	E 12
為了促進更豐富的文化多樣性，政府應該強制企業設計AI作品時投注資源在文化上。	33% 46% 19% (81)	90% 9% 0% (11)	20% 40% 40% (25)	40% 36% 22% (22)	8% 8% 8% (12)	10% 41% 0% (11)
我認為政府不應對AI進行任何審查。	12% 22% 42% (78)	61% 42% 0% (14)	5% 46% 21% (18)	4% 50% 0% (2)	7% 64% 7% (12)	4% 20% 30% (4)
為了促進技術的平等，我認為AI公司應該定期揭露資訊資料。	43% 34% 22% (81)	90% 0% 10% (10)	41% 20% 37% (24)	61% 9% 28% (21)	9% 92% 7% (14)	2% 66% 8% (2)
為了促進技術的平等，我為AI公司定期揭露資訊資料。	43% 34% 21% (82)	90% 0% 10% (10)	42% 19% 38% (24)	59% 14% 22% (22)	0% 100% 0% (4)	3% 50% 20% (10)
我認為，是對AI公司自己該負責任和生産的內容有問題。	52% 22% 25% (71)	100% 0% 0% (8)	40% 10% 45% (20)	32% 28% 16% (21)	16% 41% 41% (12)	6% 33% 0% (9)
我認為，使用AI生成的内容，是使用他人創作的內容。	68% 20% 12% (77)	95% 10% 0% (10)	68% 9% 22% (22)	45% 36% 16% (22)	61% 38% 0% (13)	5% 0% 10% (1)
為了促進平等、應急，應該讓AI開發者定期揭露資訊資料。	28% 50% 21% (71)	37% 37% 25% (8)	20% 40% 35% (20)	18% 77% 4% (22)	38% 4% 18% (11)	40% 30% 30% (10)
我認為AI應該尊重個人隱私，並定期揭露資訊資料。	55% 28% 16% (80)	60% 30% 10% (10)	48% 28% 24% (25)	81% 9% 9% (22)	50% 53% 16% (12)	18% 63% 18% (11)
在生成內容時，如果有任何疑慮，應該馬上停止。	46% 23% 29% (71)	88% 11% 0% (8)	42% 10% 47% (18)	81% 14% 22% (21)	16% 41% 41% (13)	20% 60% 20% (10)
台灣應該要對AI產生的內容採取法律監視，不要一昧地相信AI。	63% 22% 14% (77)	50% 40% 10% (10)	63% 13% 22% (22)	71% 14% 14% (21)	71% 21% 7% (14)	50% 40% 10% (10)
開發者應該為AI產生的內容負責。	53% 27% 18% (78)	90% 10% 0% (10)	47% 26% 26% (23)	63% 9% 27% (22)	50% 11% 11% (9)	11% 77% 11% (9)
AI開發者必須為AI所產生的內容過濾，避免資訊傳播錯誤。	41% 27% 31% (74)	68% 9% 33% (8)	36% 18% 45% (22)	68% 18% 13% (22)	9% 45% 45% (21)	8% 50% 20% (10)
我認為，AI開發者應該基於公理，維護和促進社會正義。	83% 5% 11% (78)	100% 0% 0% (10)	77% 4% 18% (22)	80% 0% 9% (22)	71% 21% 7% (14)	80% 0% 20% (10)
AI必須不斷地公開方法論和資料庫，以增進數力。	19% 48% 31% (69)	20% 42% 12% (8)	16% 38% 44% (18)	5% 70% 25% (20)	25% 41% 5% (12)	50% 12% 37% (8)
AI生成的內容不能沒有著作權。	36% 36% 24% (70)	54% 45% 0% (1)	26% 25% 40% (20)	47% 28% 23% (21)	9% 63% 27% (11)	30% 50% 20% (10)
AI開發者應該向公眾說明其道德原則，並應盡道德責任。例如：因為AI人力勞動的勞工，需要政府保護及規範。	48% 33% 17% (74)	70% 20% 10% (10)	67% 19% 23% (21)	47% 33% 19% (21)	25% 50% 25% (12)	40% 60% 0% (8)
AI開發者應該與民間團體接觸，並定期交換意見。	29% 60% 10% (88)	80% 10% 0% (10)	17% 53% 28% (28)	89% 91% 0% (28)	10% 64% 6% (18)	72% 27% 0% (10)
台灣對AI的規範與國際標準一致。	40% 26% 33% (72)	100% 0% 0% (8)	28% 21% 52% (29)	45% 25% 30% (20)	8% 50% 41% (12)	5% 53% 11% (9)
因為AI的規範與國際標準一致，所以我們希望它能符合國際標準，並在全世界範圍內被廣泛採用。	55% 27% 17% (78)	100% 2% 0% (10)	50% 13% 36% (22)	80% 4% 14% (21)	30% 61% 7% (13)	0% 50% 10% (10)



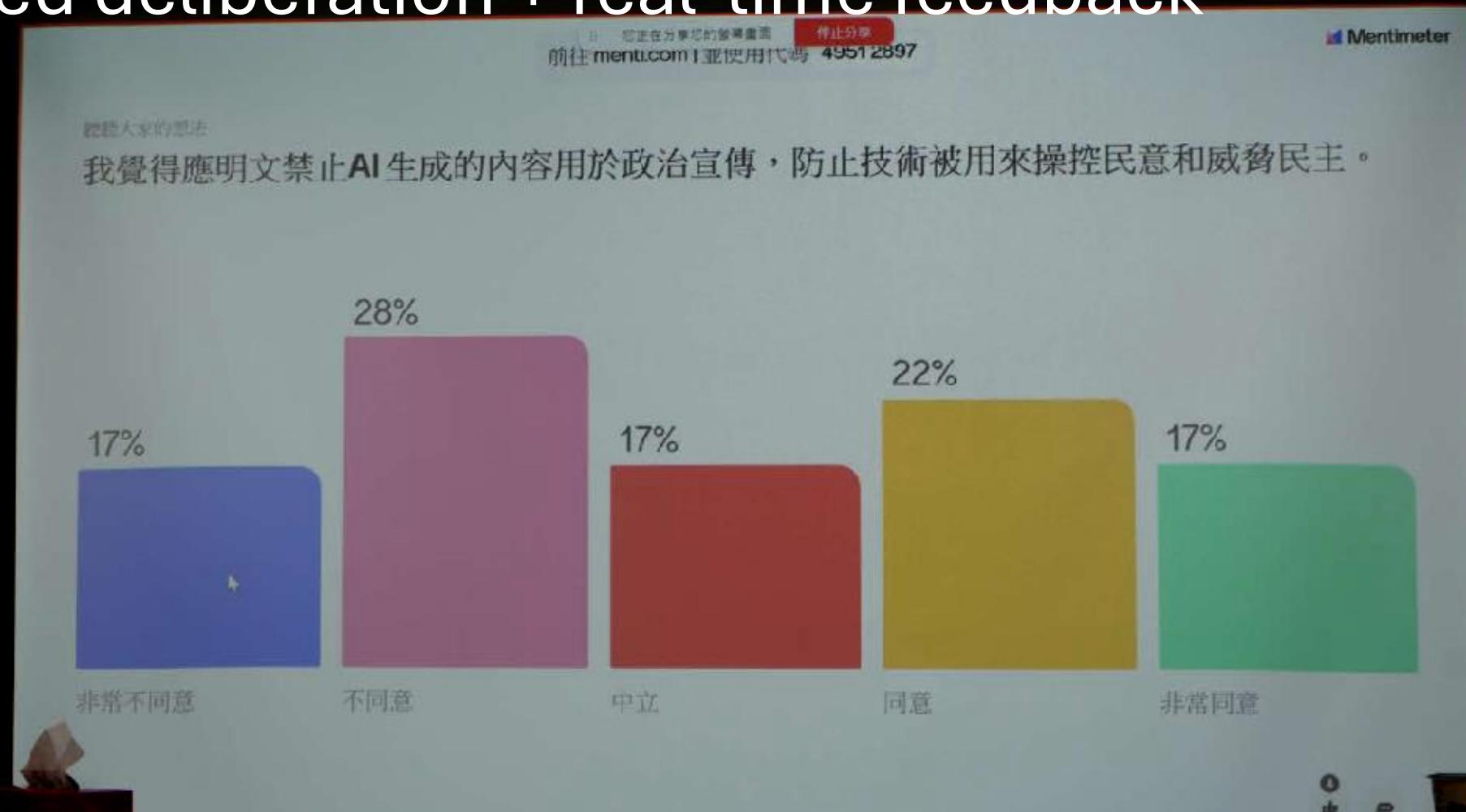
Online voting
Board Participation

Consensus Workshop
Identifying consensus and division

Solution Workshop
Brainstorm actionable solutions

Method: ReadTheRoom Deliberation

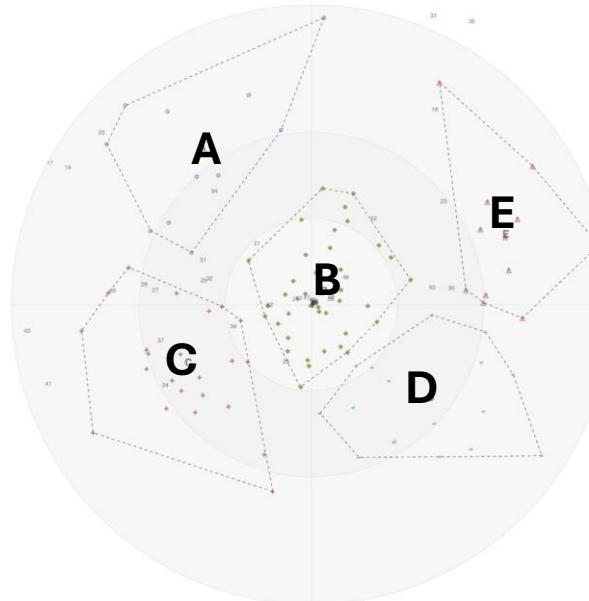
Structured deliberation + real-time feedback



Voting-guided Deliberation

Online: First, identify the issues that divide the groups

STATEMENT	OVERALL 105	A 11	B 43	C 23	D 16	E 12
17 我認為現在的中文太僵硬了，政府應該讓它更流動、更自然一些。	33% 46% 19% (81)	90% 9% 0% (11)	20% 40% 40% (20)	40% 36% 22% (22)	8% 63% 8% (15)	18% 81% 0% (11)
18 我認為政府不應對AI進行任何審查。	15% 26% 12% (73)	40% 40% 0% (10)	5% 66% 27% (18)	4% 95% 0% (22)	7% 84% 7% (15)	40% 30% 30% (10)
19 計算AI的公司應該公開資訊與資料。	43% 34% 22% (81)	90% 0% 10% (10)	41% 20% 37% (24)	61% 9% 28% (21)	0% 92% 7% (14)	25% 66% 8% (12)
20 計算AI的公司應該公開資訊與資料。	43% 34% 21% (82)	90% 0% 10% (10)	42% 19% 38% (20)	59% 18% 22% (22)	0% 100% 0% (14)	30% 50% 20% (10)
21 我認爲，是好的AI公司自己應該負起道德生產的責任，是壞的AI公司自己應該負起道德生產的責任。	52% 22% 25% (77)	100% 0% 0% (9)	45% 10% 45% (20)	52% 28% 10% (21)	16% 41% 41% (12)	66% 33% 0% (9)
22 我認爲，使用AI生產的內容，是應用者自己應該負責的。	66% 20% 12% (77)	90% 10% 0% (10)	66% 9% 22% (22)	45% 36% 16% (22)	61% 38% 0% (15)	90% 0% 10% (10)
23 為了提升效率，應該最大化地利用好的AI和變更AI的規範，並強化AI的規範。	28% 50% 21% (77)	37% 37% 25% (8)	25% 40% 30% (20)	18% 77% 4% (20)	36% 45% 18% (11)	40% 30% 10% (10)
24 高於道德標準，才能保證AI的正確性，如果要達到道德標準，必須要強化AI的規範。	55% 28% 16% (80)	60% 30% 10% (10)	49% 28% 24% (20)	81% 9% 9% (20)	50% 50% 16% (12)	18% 42% 18% (11)
25 在內容方面，如果要達到道德標準，必須要強化AI的規範。	48% 23% 29% (77)	88% 11% 0% (9)	42% 10% 47% (19)	61% 14% 23% (21)	18% 41% 41% (12)	20% 60% 20% (10)
26 自然語言處理有已經達到人類水準的工具替換人，不是一昧地追求機器標準。	63% 22% 14% (77)	50% 40% 10% (10)	63% 13% 22% (22)	71% 14% 14% (21)	71% 16% 7% (15)	50% 40% 10% (10)
27 請從道德和經濟兩方面考慮的內容。	53% 27% 18% (78)	80% 10% 0% (10)	47% 26% 26% (20)	63% 9% 27% (20)	50% 41% 5% (12)	11% 72% 11% (8)
28 AI應該必須滿足人所產生的內容需求，但不能影響道德標準。	41% 27% 31% (74)	60% 2% 33% (8)	36% 18% 41% (20)	68% 18% 15% (22)	9% 42% 4% (11)	10% 70% 20% (10)
29 強調AI，AI應該基於公眾、使用者的權利和原則，避免損害用戶的個人隱私。	83% 56% 11% (78)	100% 0% 0% (10)	77% 4% 18% (20)	90% 0% 9% (22)	71% 21% 7% (15)	80% 0% 20% (10)
30 AI應該不能因為公眾利益和經濟利益，以道德為權力。	19% 48% 51% (68)	25% 62% 12% (8)	16% 38% 44% (16)	5% 70% 25% (20)	25% 41% 33% (12)	50% 12% 37% (8)
31 AI內容的內容不應該有作用。	36% 38% 24% (73)	54% 45% 0% (11)	30% 25% 40% (20)	47% 28% 23% (21)	9% 63% 27% (11)	30% 50% 20% (10)
AI應該要能滿足外在的需求，應由政府負責監督，例如：協助人際關係、促進社會平等、推動社會工程、促進社會福利政策和規範。	48% 35% 17% (74)	70% 20% 10% (10)	37% 19% 23% (20)	47% 33% 16% (17)	25% 50% 25% (12)	40% 40% 0% (10)
AI的生產過程應該盡量交由市場解決，政府不應該介入。	29% 60% 10% (85)	80% 10% 10% (10)	17% 53% 28% (20)	8% 91% 0% (23)	15% 84% 0% (12)	72% 27% 0% (11)
34 自然語言AI的應用範圍應廣泛一致。	40% 28% 33% (72)	100% 0% 0% (8)	26% 21% 52% (20)	45% 25% 30% (20)	8% 50% 41% (12)	55% 35% 11% (8)
因為AI帶來的影響很大，政府對於AI的特點應有清晰的辨識，並強化規範的制定，要用更嚴格的法律來規範。	58% 27% 17% (76)	100% 0% 0% (10)	50% 13% 36% (22)	80% 4% 14% (21)	30% 61% 7% (13)	0% 90% 10% (10)

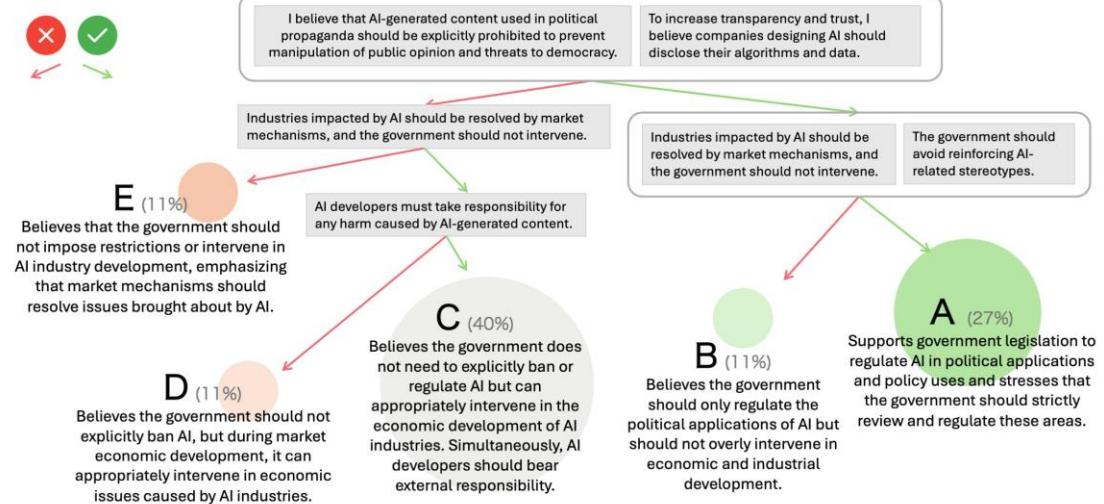


Wikisurvey data

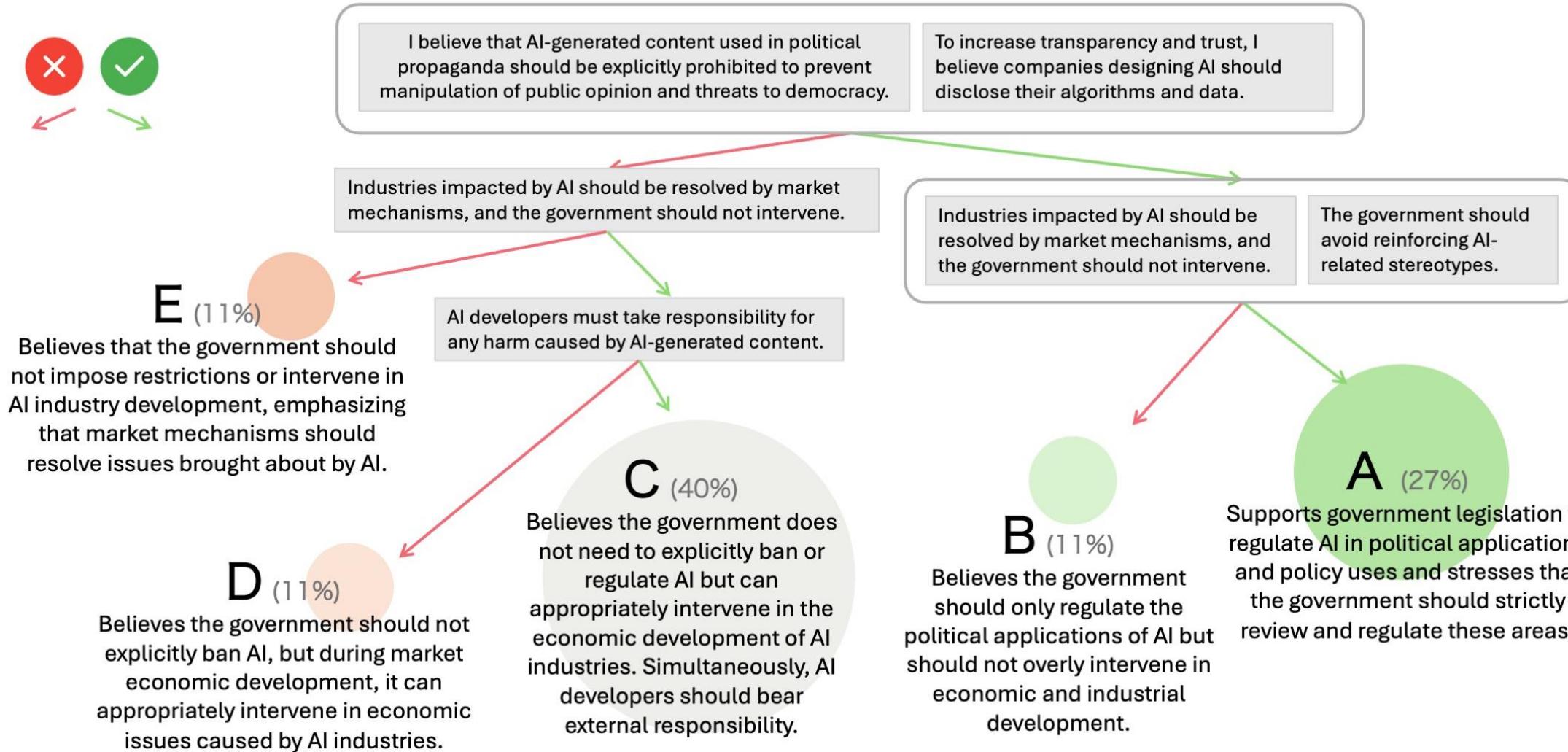
Opinion groups

Find key statements
that divide the groups

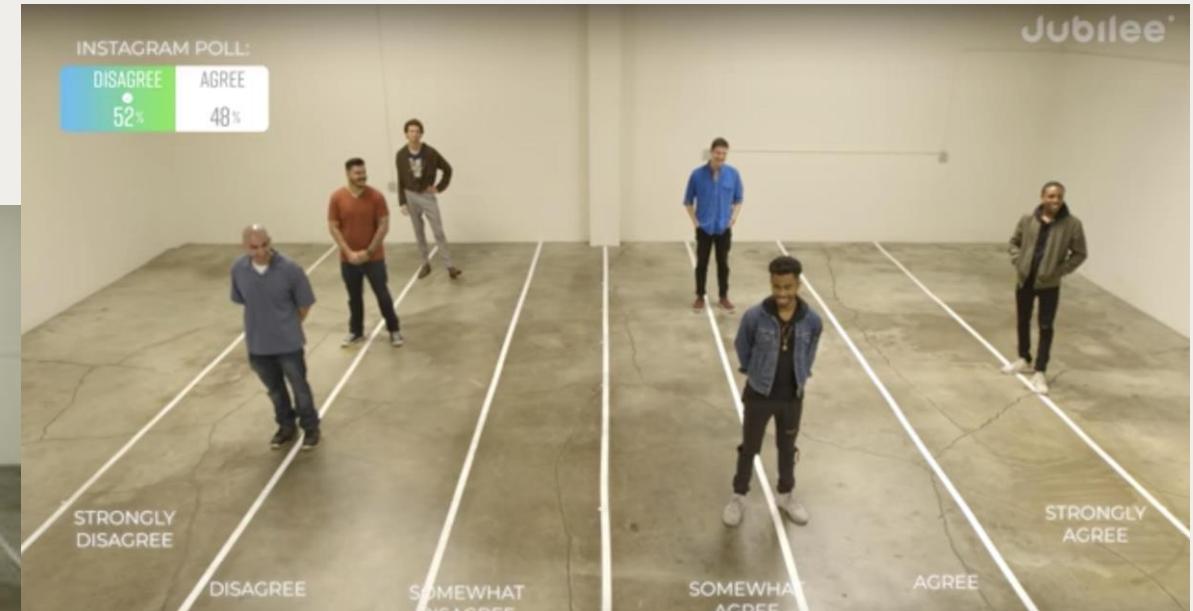
vTaiwan Decision Tree: Finding Dividing Points, Not Just Looking at Majority Opinions



Participants learn who they (try to) represent online



Visualising the Opinion Landscape for Mutual Understanding



ReadTheRoom: Visualise the Opinion Landscape

Using real-time online voting to gamify openmindedness

1. Vote: before
deliberation

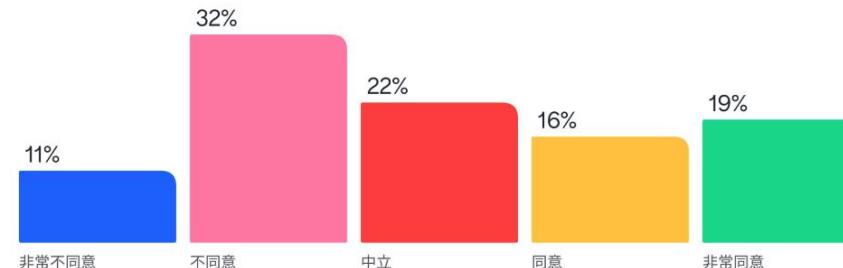
3. Vote again:
after deliberation

2. Deliberate: Each
camp sharing why they
hold this opinion

4. Deliberate again:
Share why you change
your opinion

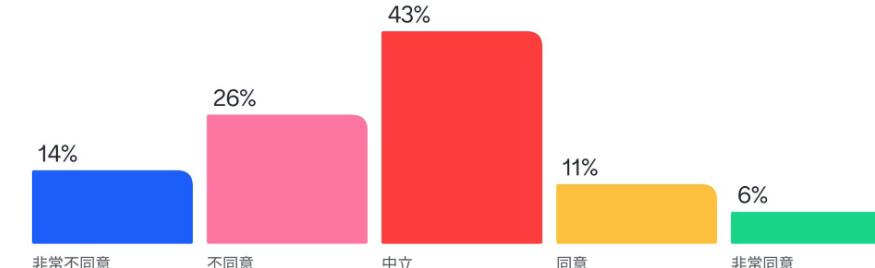
聽聽大家的想法

為了增加透明度和信任度，我認為設計AI的公司應該公開其算法與資料。



聽完大家分享，你的想法有什麼改變嗎？為什麼呢？

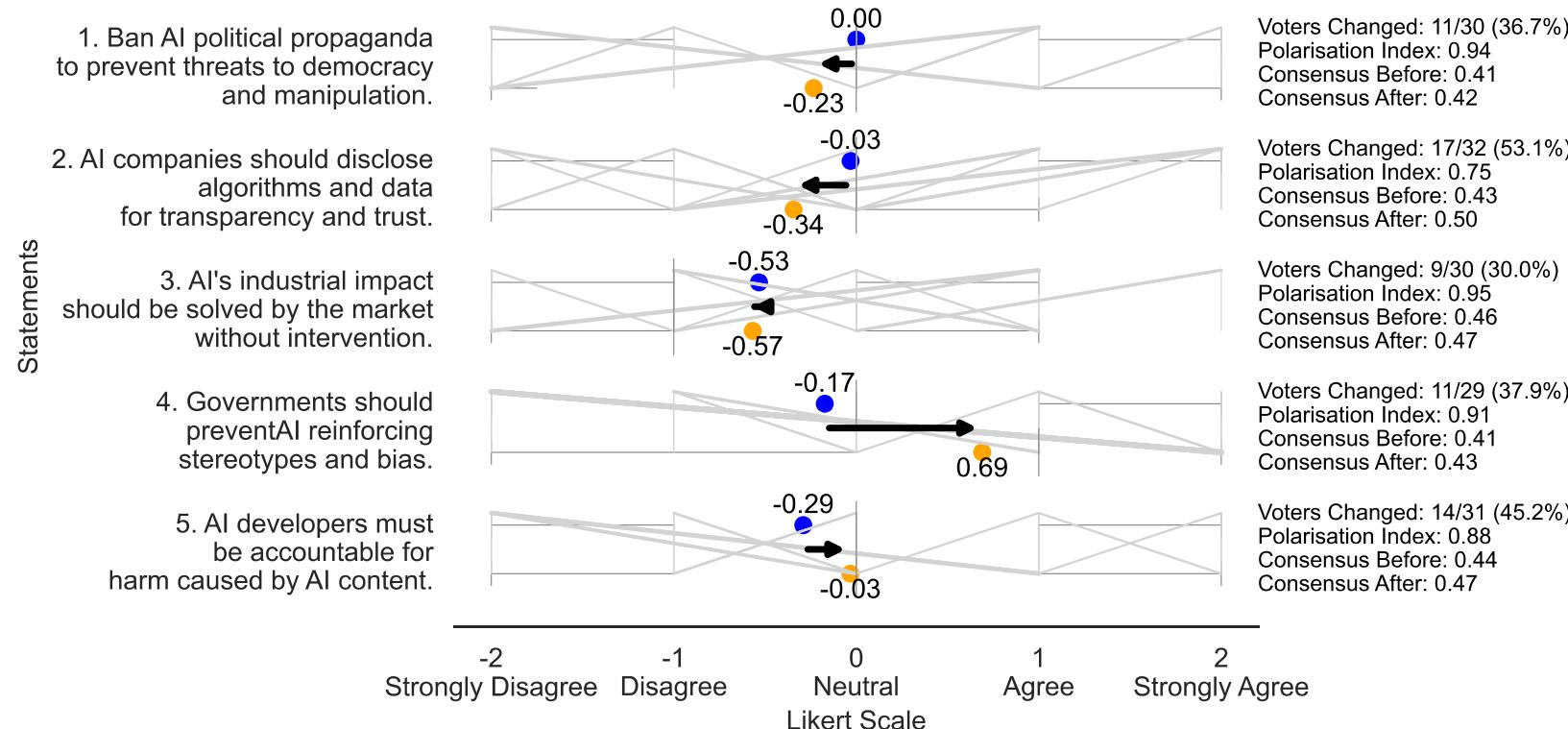
為了增加透明度和信任度，我認為設計AI的公司應該公開其算法與資料。



Spectrum-based voting showed preference shifts pre- and post-deliberation

ReadTheRoom: Voting-guided Deliberation

Giving people a second chance after deliberation reduces polarisation and increases consensus

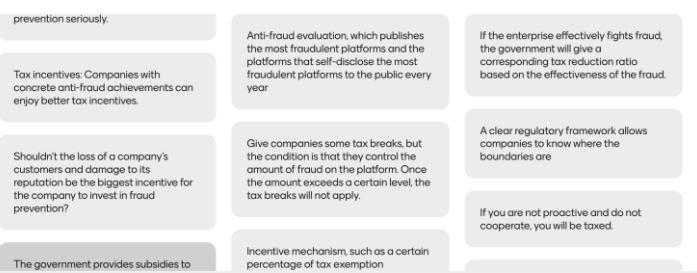


Qualitative survey shows that this method
supports reflection and learning

Solution Workshop: AI-in-the-loop Deliberation

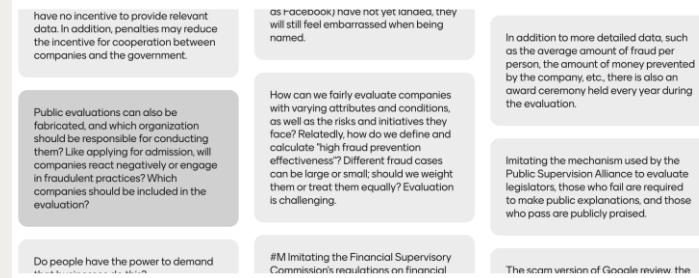
Propose solution ideas

如何可以增加企業投入打詐的誘因？我們來提出實際的行動方案

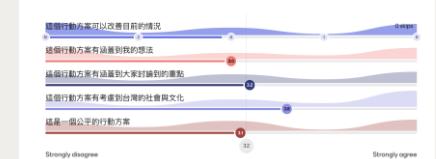


Develop Action Plan

集思廣益，關於這個行動方案：「反詐表現透明化與評比機制。建立年度企業防詐評比制度，公開揭露各平台的詐騙數據與治理成效，提升企業對品牌聲譽的重視與投入誘因」，有哪些可以更完善的地方？

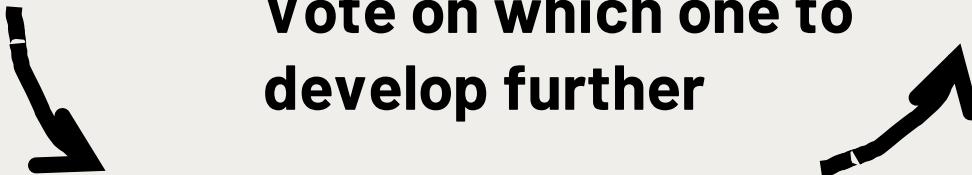


你對大家討論出來的這個行動方案有什麼想法？

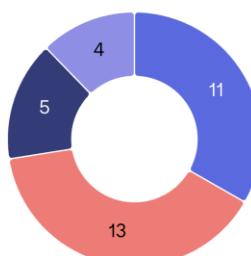


Rate final plan

Vote on which one to develop further



Now we have sorted out four action plans. Let's vote to decide which one to discuss in depth.



LLM generates solution directions

- 11 建立社區防詐網絡，結合老人共餐、社大課程與在地組織，定期舉辦經驗分享、識詐教學與實境模擬，讓防詐教育融入日常生活。
- 13 推出全民參與獎勵機制，設計像「你好棒獎狀」、檢舉獎金、減稅優惠等獎勵措施，讓民眾在生活中對抗詐騙有成就感與參與感。
- 5 舉辦開放徵案與反詐騙創意競賽，例如黑客松、詐騙劇本徵稿或實境秀，鼓勵民間提出創新點子並由政府支持落地實行。
- 4 納入義務教育及社區宣導，由公部門深入學校與社區，強化資訊識別力與社會互助意識，並建立匿名分享平台協助曾受害者互助復原。



LLM summarises into action plan

行動方案#1 意見整理：「反詐表現透明化與評比機制。」

可建立由政府委託的中立第三方負責的「平台防詐評比制度」，針對社群、通訊、電商、拍賣、銀行等不同類型平台，依據規模、風險屬性設計差異化指標，例如：被害人數、總損失、人均損失、詐騙金額、資料揭露度等，搭配加權評分。企業可自評填報，但必須附上佐證資料，並與165數據與警政資料交叉比對，以降低黑數與造假風險。

境外企業雖無法可管，仍能納入評比名單並公布其在台風險指數，設計類似「詐騙版Google review」，讓用戶搜尋或連結時即顯示風險警示，例如「2025年受害人數：3,200人、損失金額：8.1億元」。雖無法強制，但國際平台在意輿論與品牌聲譽，仍會受到壓力。

評比結果每年公布，表現良好者公開表揚、頒發標章，並可獲稅務或宣傳資源獎勵；表現墊底者須公開說明並限期改善，連續不改善者列入高風險名單，限制公部門合作與廣告推播。大型平台重罰壓力、小型企業則給予鼓勵。廣告強制揭露制度也可參考金管會做法，要求平台在廣告或首頁揭示年度詐騙數據，不得隱藏。例如：「本平台2025年被害人數達3,200人，總損失8億元」，用標準語速、字級與時長呈現，讓使用者一打開就知到。

為擴大全民參與並補足資料缺口，可加碼檢舉車手與詐騙管道的獎金，讓所有民眾都能參與通報與監督。整體制度設計原則是：由上而下施壓、由下而上補洞、由內而外推動透明，讓企業無法迴避。

Takeaway

Voting can do more than just aggregate preferences—with digital support, they can be designed to enhance citizens' capacity for informed participation.



Other common democratic innovations



- **Online participation platforms**
(Decidim, Consul, Your Priorities)
- **Deliberation tools** (Polis, Loomio)
- **Liquid democracy platforms**
(LiquidFeedback)
- **E-voting and digital identification**
- **AI assisted participation**
(summarisation, moderation, proposal support)

Voting Advice Application: Smartvote/ Palumba

94

smartvote 

English ▾ Share Login

Einwohnerratswahlen Brugg 2025 ▾

Start Voting advice Database

1. Sozialstaat, Familie & Gesundheit (0/4) ▾

1. Soll die Stadt den gemeinnützigen Wohnungsbau verstärkt fördern (z.B. finanzielle Unterstützung von Wohnbaugenossenschaften, Zurverfügungstellung von vergünstigtem Bauland)?

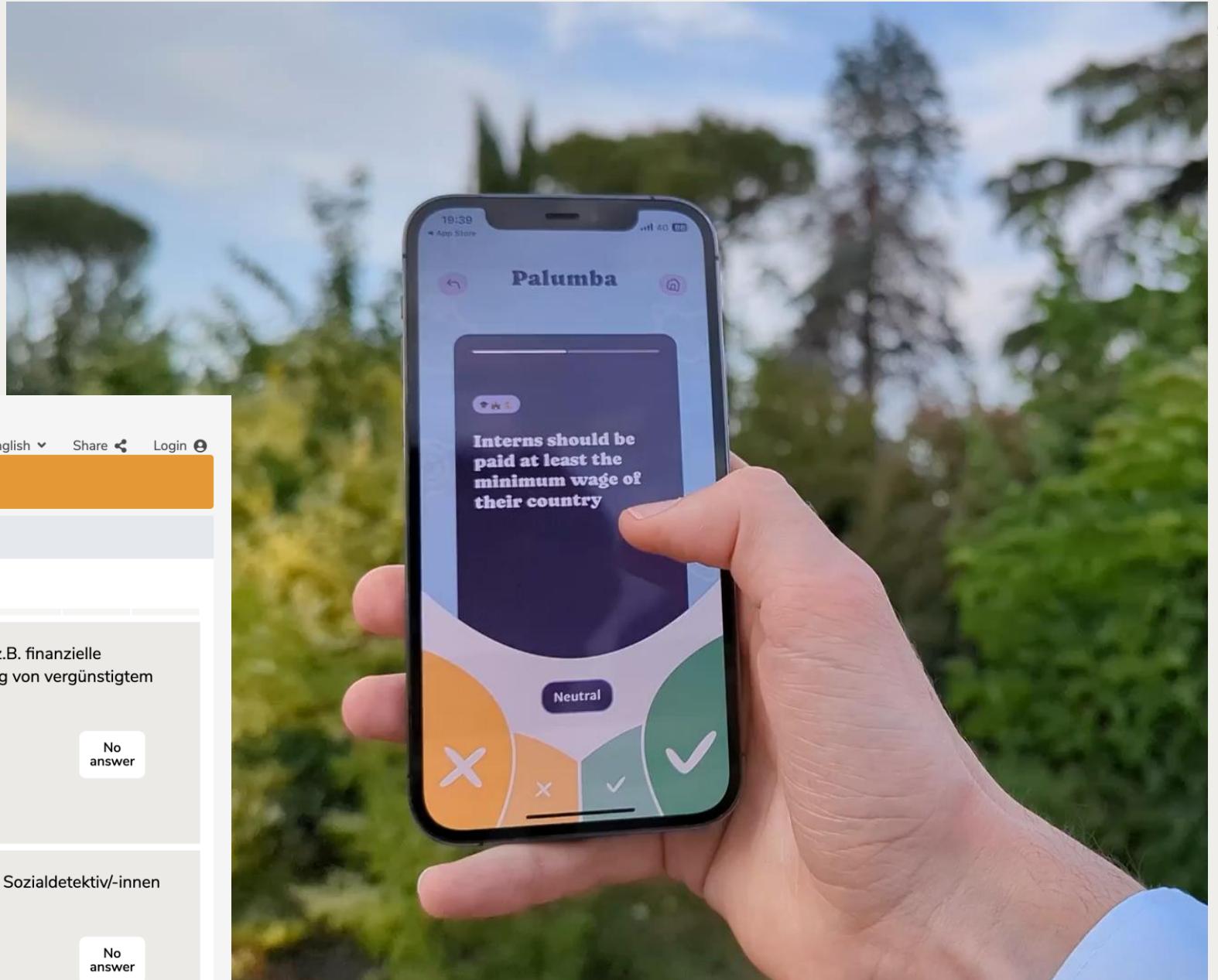
Yes Rather yes Rather no No
No answer

Weight answer: - = +

2. Soll die Stadt Brugg bei Verdacht auf Sozialhilfemissbrauch vermehrt Sozialdetektiv/-innen einsetzen?

Yes Rather yes Rather no No
No answer

Weight answer: - = +



みんなの意見: 経済・税制・雇用

あなたの意見をタップすると、あなたのアイコンが世論地図上でリアルタイムに動きます

「経済・税制・雇用」って?(政策比較ページへ)

企業の賃上げを促すため、法人税の優遇や減税措置をさらに拡大すべきだ

賛成

どちらでもない

反対

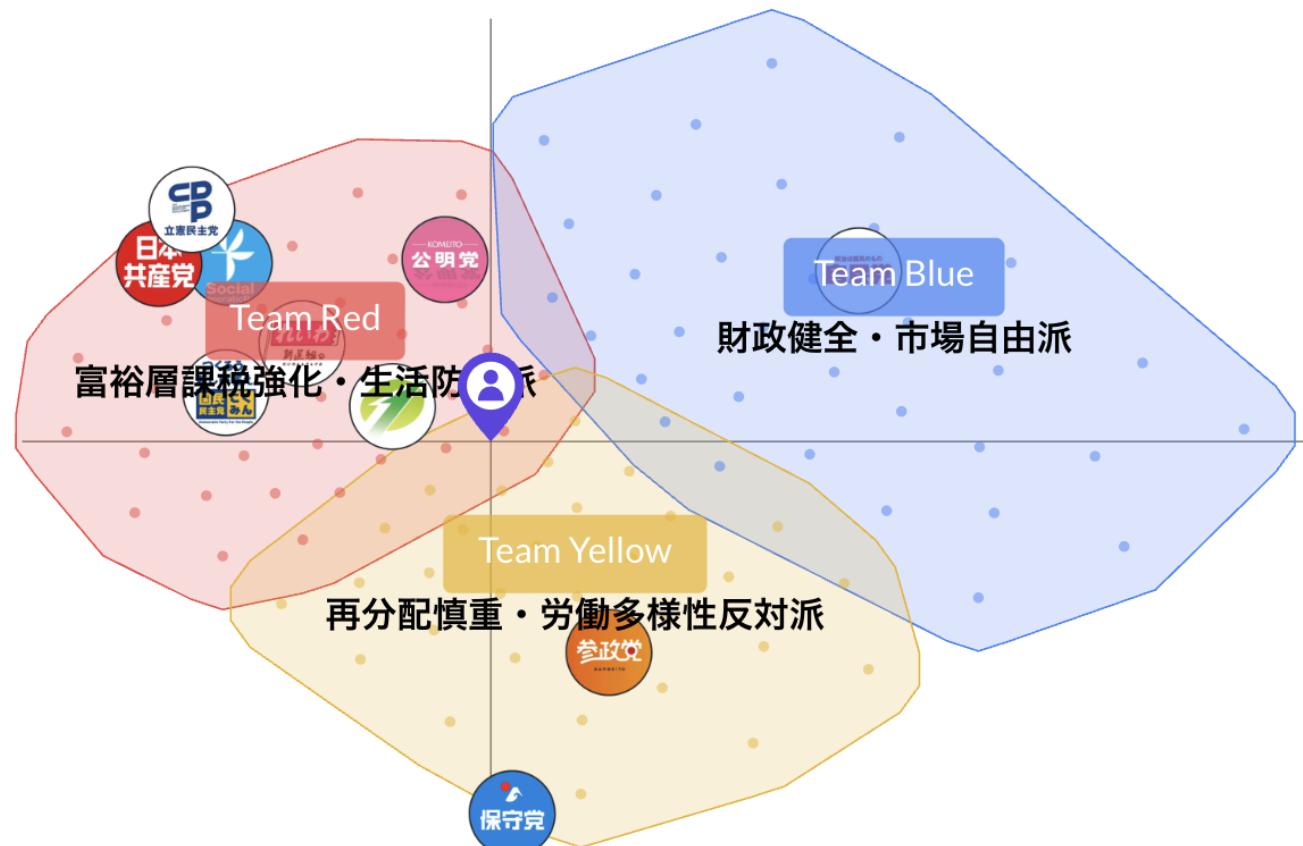
わからないのでこの設問はスキップ

意見の可視化結果を見る



Political Information Platform: Japan Choice

意見の可視化



Risks and design challenges



1. Inequality

Digital divide, participation by already privileged groups.

2. Power and control

Platforms controlled by governments or vendors, risk of manipulation and agenda setting.

3. Quality of deliberation

Echo chambers, trolling, shallow engagement, need for facilitation.

4. Security and trust

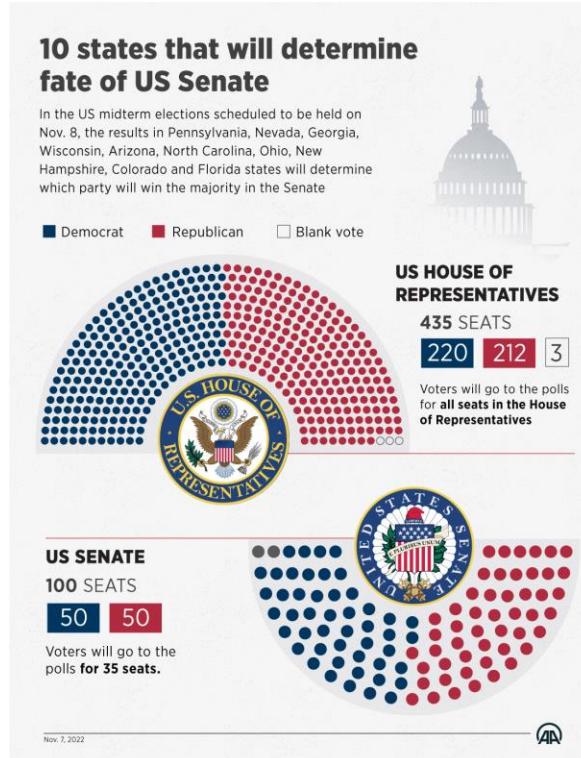
Concerns about data protection, identity, hacking, deepfakes.



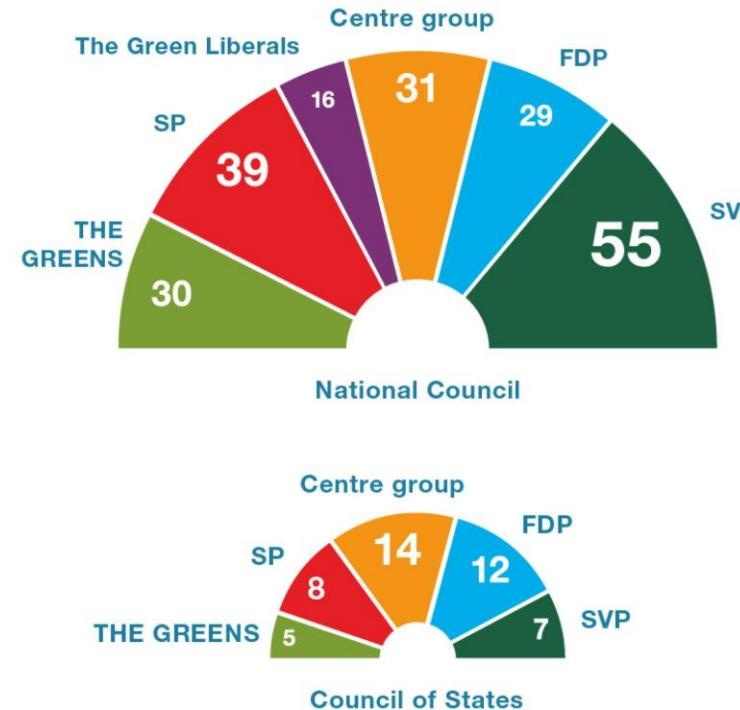
From Social Choice to AI Alignment

Joshua C. Yang / joyang@ethz.ch

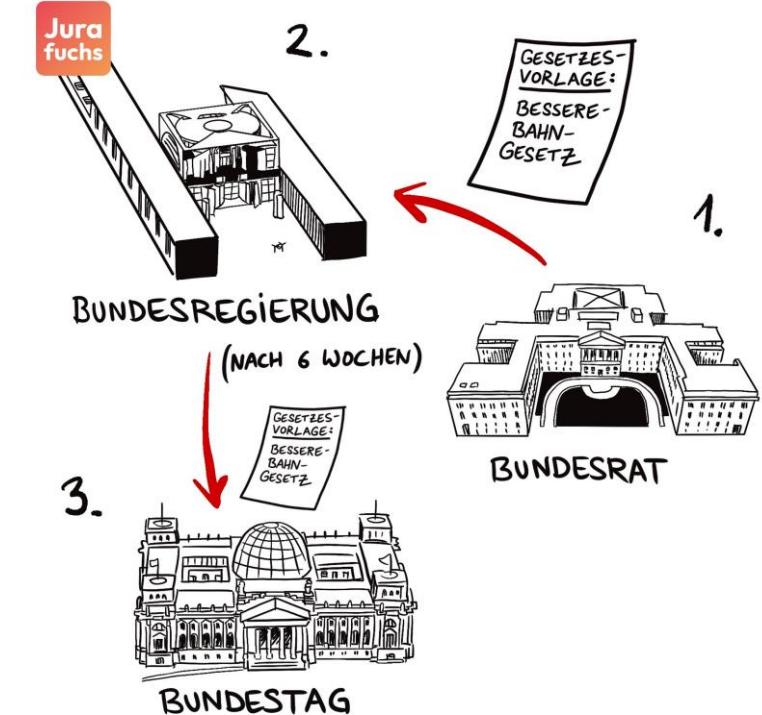
Q: Why do democracies have bicameralism?



USA

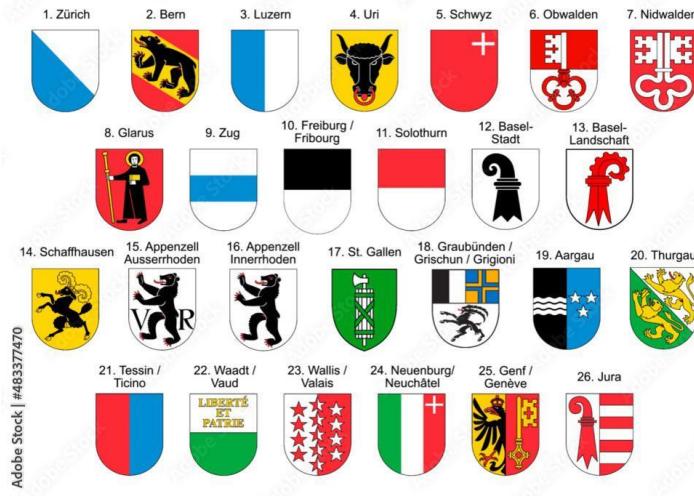


Switzerland

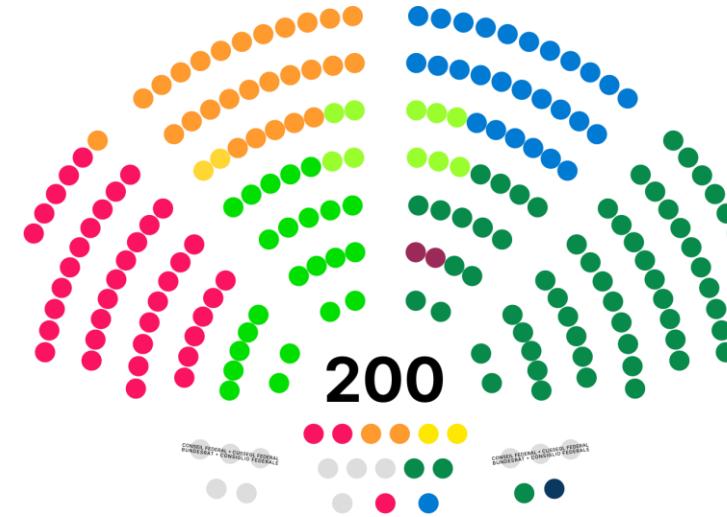


Germany

Localism vs. Proportional Representation



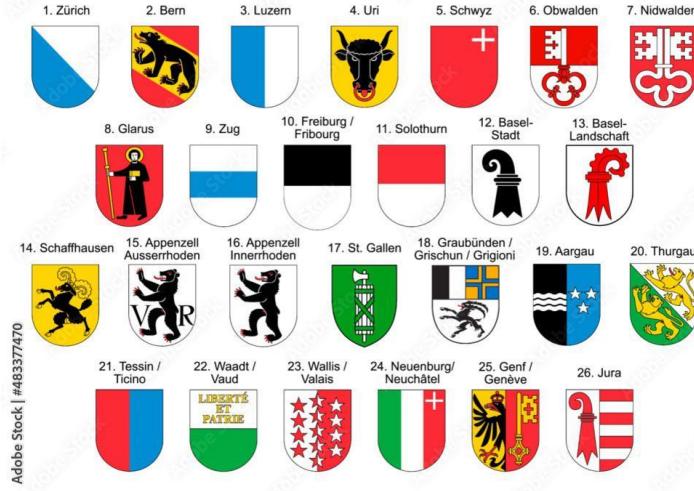
Community based



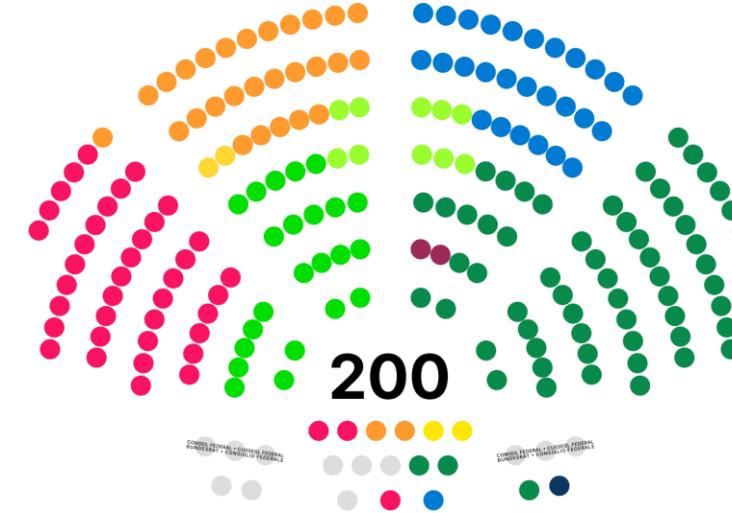
Population based

← US CH DE DK NL →

Localism vs. Proportional Representation



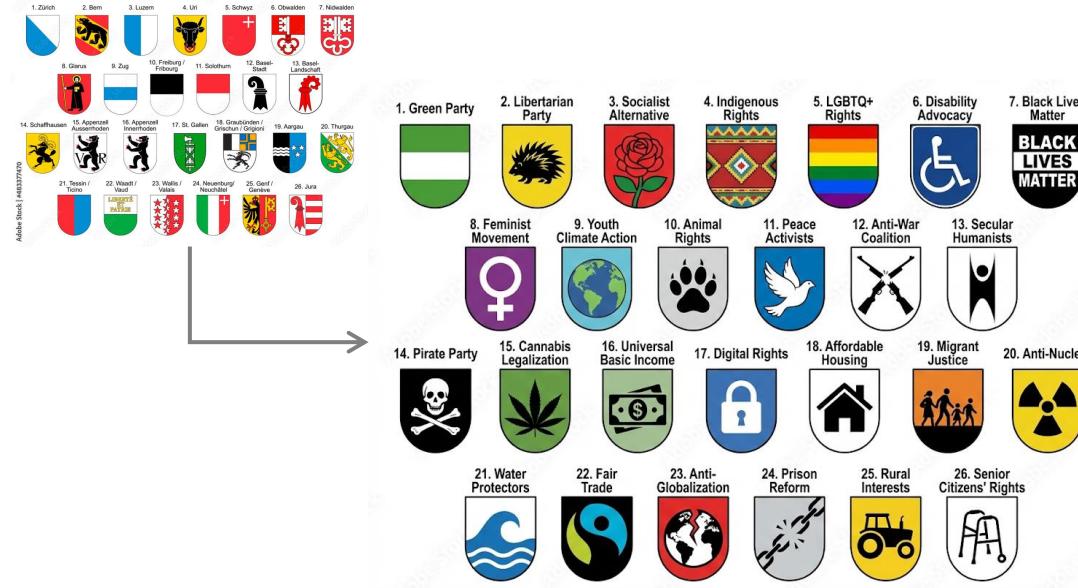
Community based



Population based

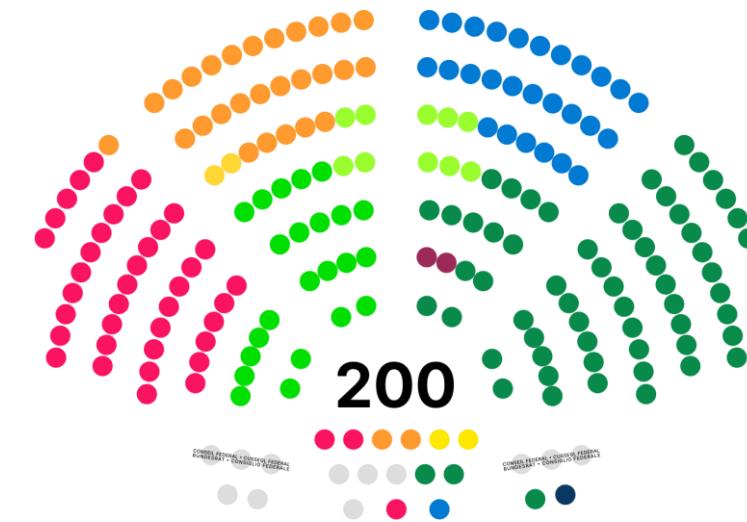
Which one is fairer?

Localism in a Globalised and Digital World



Community based

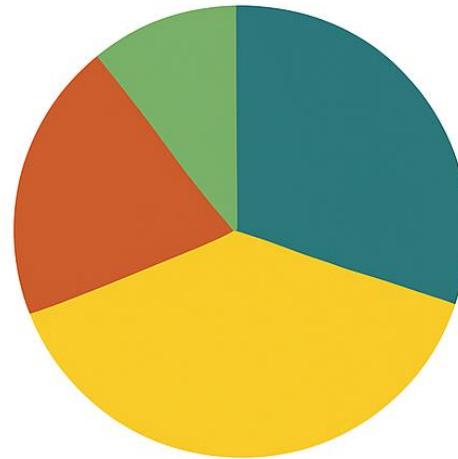
Geographical communities
→ Issue-based digital communities



Population based

Proportional Representation vs. Winner-take-all

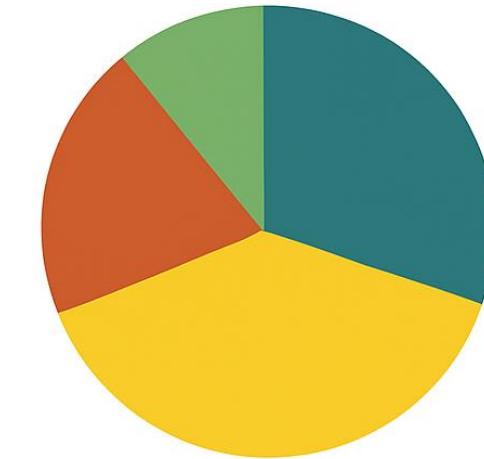
Voter Support



Winner-Take-All



Proportional Representation



How can we design Proportional Representation into our systems?

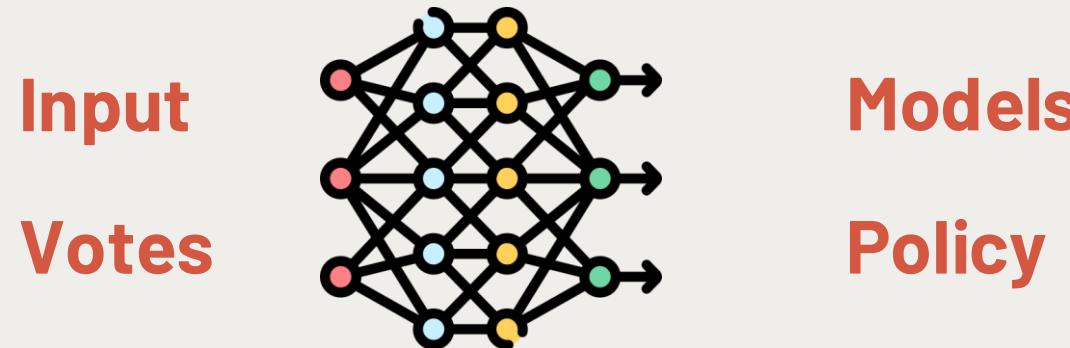
Representation and Aggregation in Democracy

How decisions are made?

Who is represented?

	Winner-take-all	Proportional Representation
Population based	<ul style="list-style-type: none"> • Single-winner elections • Presidential elections • Referendum • Deciding what to have for dinner 	<ul style="list-style-type: none"> • Multi-winner elections • Bundestag • Participatory budgeting • Playlist ordering based on user ratings
Community based	<ul style="list-style-type: none"> • UN votes/International organisations where unanimity is required • Worker's Union strike • Indigenous veto rights 	<ul style="list-style-type: none"> • University committees with fixed student, faculty, staff seats • Public broadcasting governance balancing language regions

Is AI the (rigged) democracy of humanity?



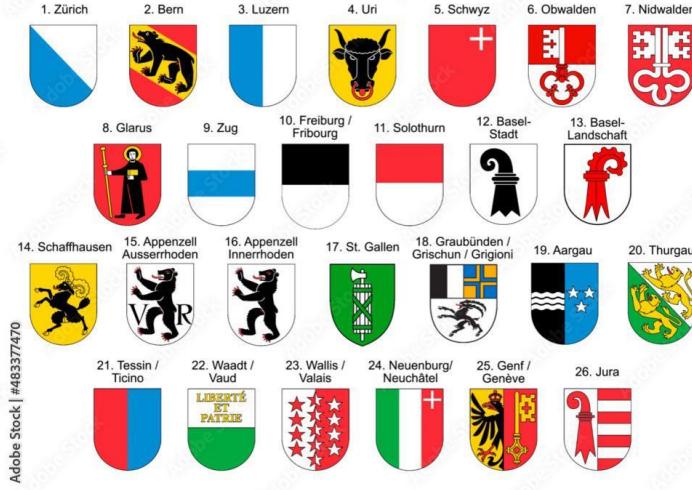
Democracy

Citizens
Votes
Ballots
Electoral system
Representatives
Election outcome
Minority voices

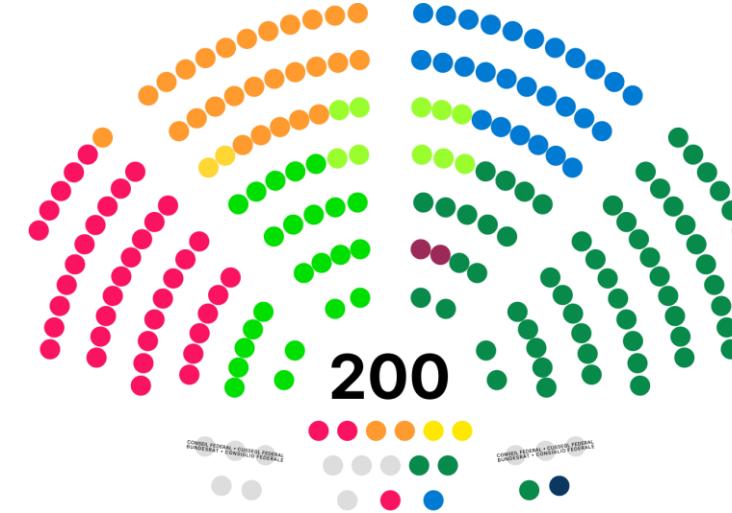
AI training

Humans whose data or feedback is used
Training signals, labels, ratings, preferences
Data points, comparisons, rankings
Training objective and aggregation method
The trained model's behaviour
The deployed model
Edge cases, marginal users, underrepresented norms

Which one are we using in AI alignment?



Community based



Population based

Representation and Aggregation in AI training

How decisions are made?

	Winner-take-all	Proportional Representaion
Who is represented?	Population based Average preference optimisation E.g. RLHF and DPO style alignment, Red-teaming focused on common failure modes	Preference distribution modelling E.g. Fairness metrics across demographic groups
Community based	Hard constraints and veto mechanisms E.g. Content moderation and refusal rules, Safety classifiers and blocklists	Stakeholder-aware power sharing E.g. Jury learning, PRISM, and representative aggregation etc.

Broader Vision for AI in Democracy

- **Direct democracy** - AI excels at helping with issue-based decisions (like infrastructure, healthcare, environment) rather than personality-driven candidate elections
- **Open data** - When governments provide accessible budget info and policy research, AI can process vast amounts of information that individual citizens can't handle alone, creating interactive research assistants
- **Bottom-up citizen initiatives** - AI could help communities develop their own policy proposals by analysing data and researching best practices

*Digital democracy is where
citizens lead and algorithms follow.*

Thank you

Joshua C. Yang
Computational Social Science
ETH Zurich

joyang@ethz.ch