Collective Control of Artificial Intelligence: A Social Choice Approach to Democratic AI Governance

Martin Lackner

University of Applied Sciences St. Pölten

16 June 2026

Publication

AI of the People, by the People, for the People: A Social Choice Approach to Collective Control of Artificial Intelligence.

Paul Anton Bachmann, Niklas Böhmer, Lukas Daniel Klausner, and Martin Lackner. In Proceedings of the *ACM Conference on Fairness, Accountability, and Transparency 2026*.

Preprint: `https://arxiv.org/abs/2605.16291`

Artificial Intelligence

A safe prediction: The topic of AI will continue to be emotionally charged.

Artificial Intelligence

A safe prediction: The topic of AI will continue to be emotionally charged.

Many substantial reasons:

- Very fast expansion into almost all aspects of life.
- Serious risks: surveillance, job displacement, loss of human autonomy, etc.
- Society — and individuals — have very little control.
- Instead: a small number of technical (often corporate) decision-makers.

Democratic control of AI

Existing work largely focuses on the macro level:
governance and institutional design *(regulatory control of AI)*.

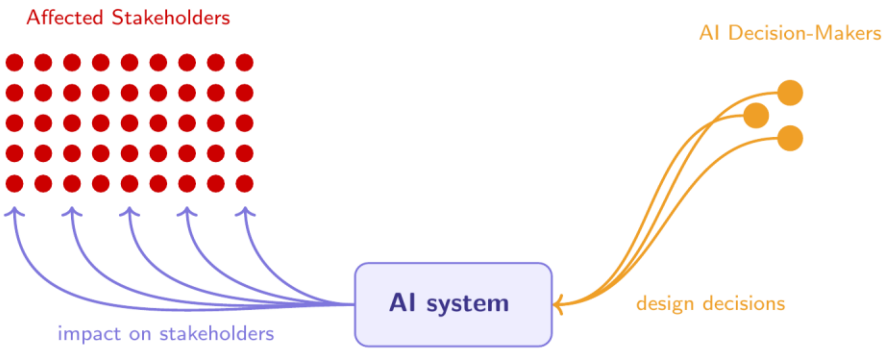
Democratic control of AI

Existing work largely focuses on the macro level:
governance and institutional design *(regulatory control of AI)*.

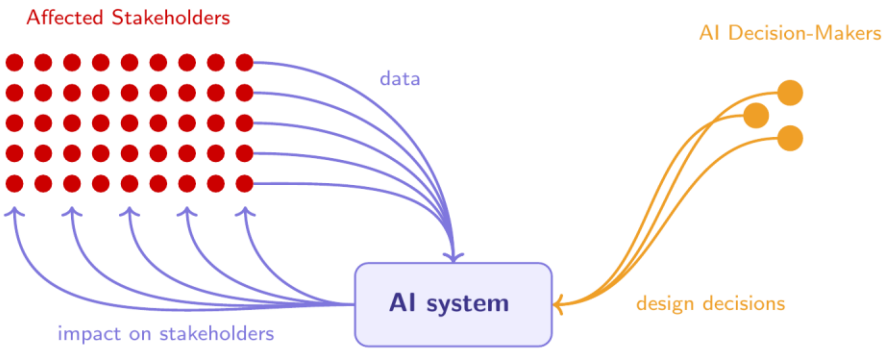
Our framework: collective control of AI
incorporating collective decision-making mechanisms directly into the AI *development and deployment* process.

- the concrete technical and socio-technical **counterpart** to regulatory control of AI

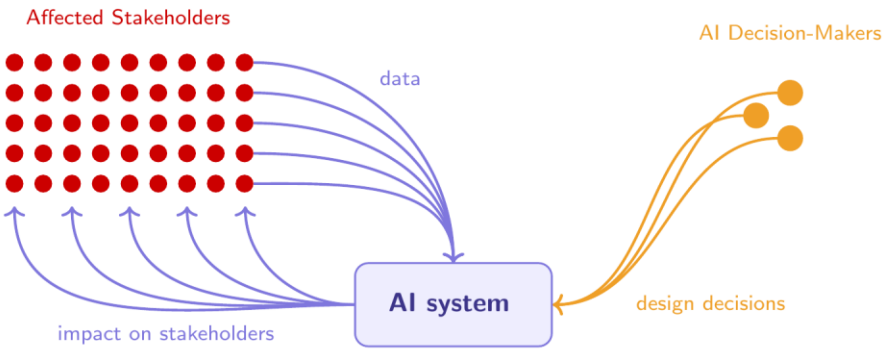
Status Quo


Affected Stakeholders
AI Decision-Makers
AI system
impact on stakeholders
design decisions

Status Quo


Affected Stakeholders
AI Decision-Makers
data
AI system
design decisions
impact on stakeholders

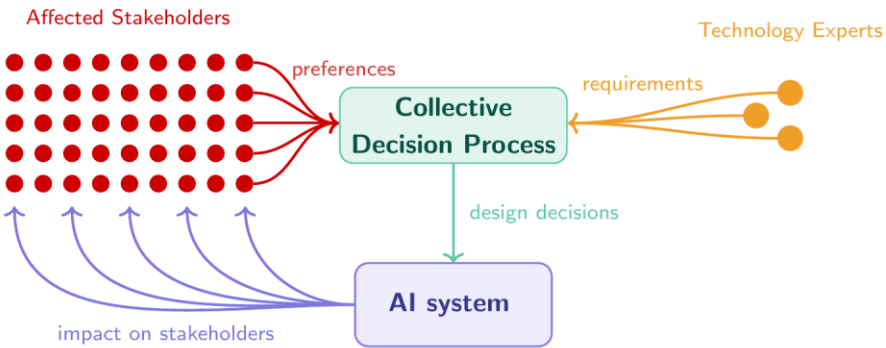
Status Quo



Goal: Collective Control of AI Systems

include affected stakeholders in AI development decisions

Collective Decisions about AI



Goal: Collective Control of AI Systems

include affected stakeholders in AI development decisions

Background: Social Choice Theory

Social Choice Theory is the formal mathematical study of collective decision problems and preference aggregation rules.

Background: Social Choice Theory

Social Choice Theory is the formal mathematical study of collective decision problems and preference aggregation rules.

Applications: decision situations with conflicting preferences, e.g. democratic elections, resource allocation, fair division, ...

Background: Social Choice Theory

Social Choice Theory is the formal mathematical study of collective decision problems and preference aggregation rules.

Applications: decision situations with conflicting preferences, e.g. democratic elections, resource allocation, fair division, ...

Central insight: Results of collective decisions heavily depend on the chosen preference aggregation rule.

Background: Social Choice Theory

Social Choice Theory is the formal mathematical study of collective decision problems and preference aggregation rules.

Applications: decision situations with conflicting preferences, e.g. democratic elections, resource allocation, fair division, ...

Central insight: Results of collective decisions heavily depend on the chosen preference aggregation rule.

Axiomatic Method

Normative evaluation of preference aggregation rules

- formulation of normative axioms (e.g. unanimity)
- study of aggregation rules w.r.t. proposed axioms

Collective Control of AI – Two Perspectives

1

What does it mean for an AI system to be aligned with collective preferences?

Collective Control of AI – Two Perspectives

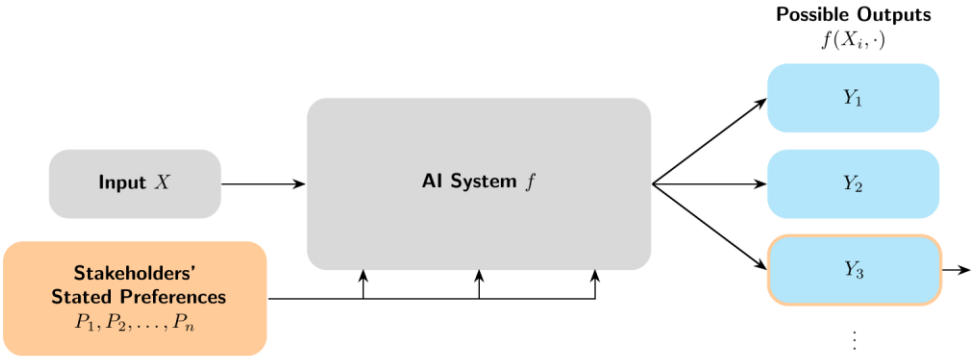
1

What does it mean for an AI system to be aligned with collective preferences?

2

How can we enable collective decision-making during AI development?

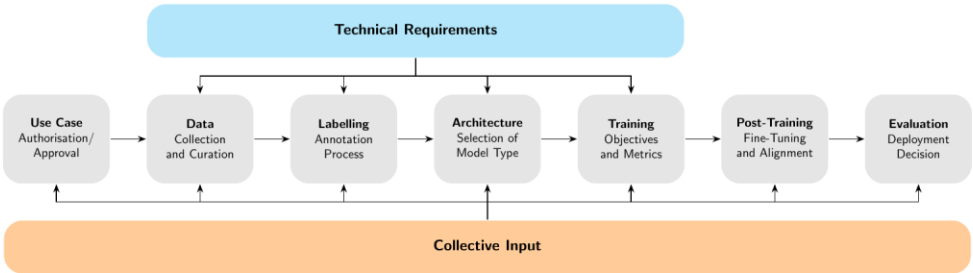
1 – AI Systems as Social Choice Functions



Social choice as lens for collective control of AI

- What *influence* can collective input have *on AI outputs*?
- How can stakeholders' preferences be *elicited* and *formalised*?
- *Which axioms* should AI systems satisfy w.r.t. the preferences?

2 – Collective Control in the ML Development Process



- How can we *enable collective decisions* in each stage?
- How can we *evaluate decisions* about the components?

We connect the decision problems in each stage to **established social choice problems**.

Open Questions and Ongoing Work

Stakeholder Identification

Who are the relevant stakeholders in each step for a given system? How can we identify them?

Open Questions and Ongoing Work

Stakeholder Identification
Who are the relevant stakeholders in each step for a given system? How can we identify them?

Preference Elicitation
How can laypeople meaningfully express their subjective preferences about AI design? How can we turn emotionally charged responses (e.g. protests) into valid collective input?

Open Questions and Ongoing Work

Stakeholder Identification
Who are the relevant stakeholders in each step for a given system? How can we identify them?

Preference Elicitation
How can laypeople meaningfully express their subjective preferences about AI design? How can we turn emotionally charged responses (e.g. protests) into valid collective input?

$\Longrightarrow$ social choice x social sciences
ongoing collaboration with Astrid Mager and Fabian Fischer (ÖAW)