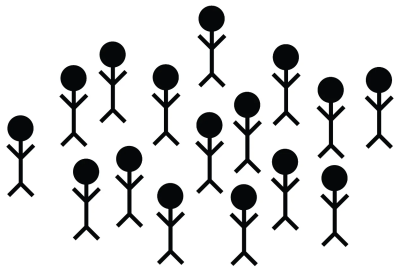
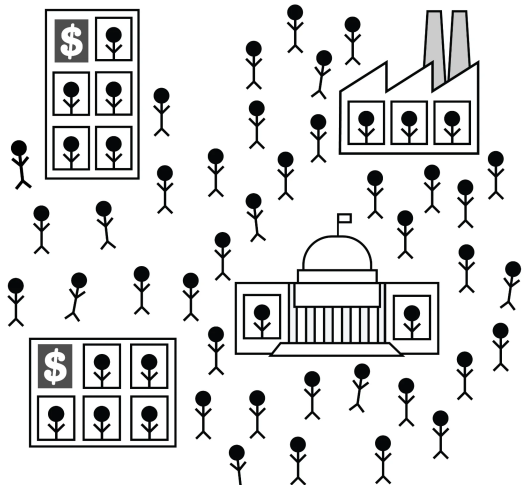
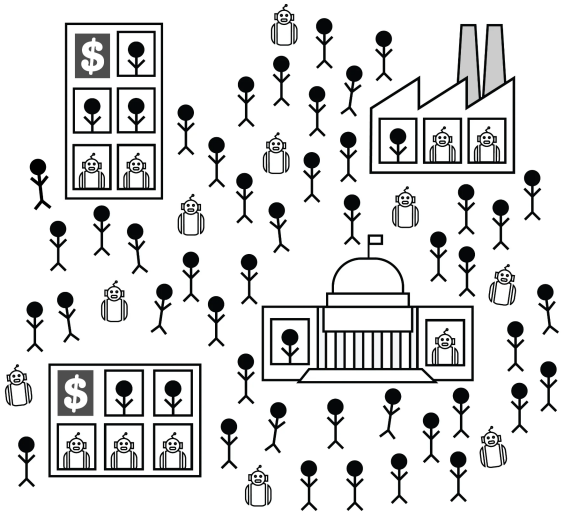
AI and the Economics of Decision-Making: Who's in Charge?

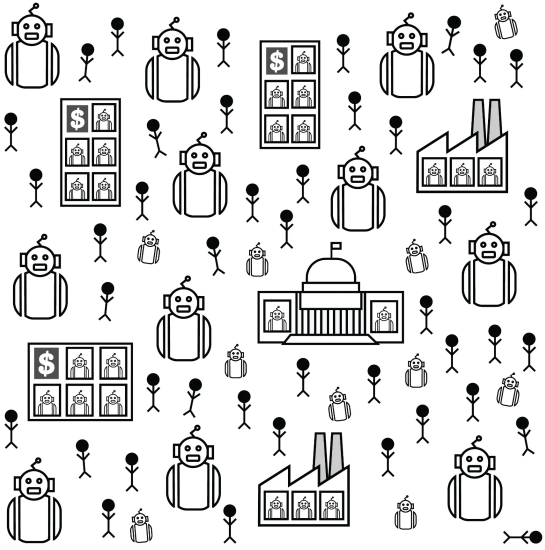
Anton Korinek
University of Virginia and EconTAI

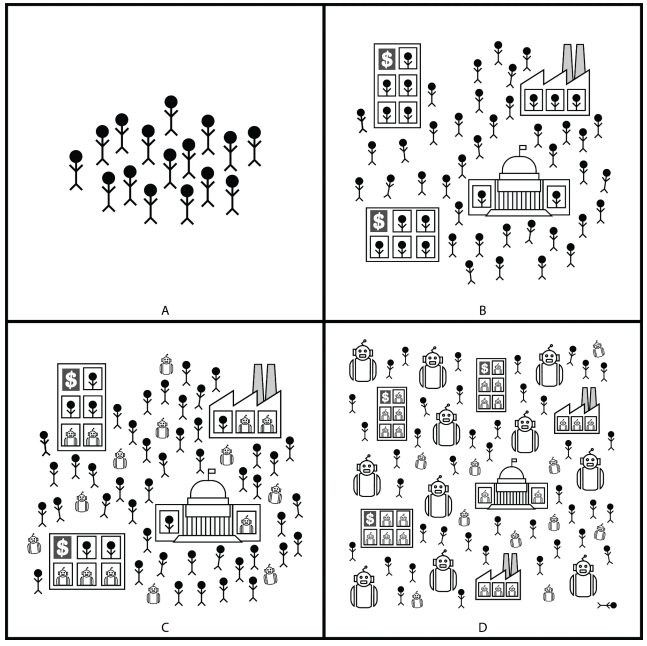
AEA 2026










A
B
C
D

AI Joins the Governance Ecosystem

The Story So Far

Era	Governance Challenge
A. Humans alone	Coordinating individuals
B. + Artificial entities (governments, corporations, markets)	Enabled massive scale—but developed "interests of their own"
C. + AI systems	New actors that evolve faster than we can govern them and increasingly operate without humans in the loop
D. + Advanced AI?	Systems beyond human comprehension and direct control

Source: Bullock, Chen, Himmelreich, Hudson, Korinek, Young & Zhang (2024), "Introduction," *Oxford Handbook of AI Governance*

The Trajectory of Economic Agency

From Tool → Advisor → Agent → Autonomous Actor

Stage	Human Role	Examples Today
AI as calculator	Human decides, AI computes	Spreadsheets, optimization
AI as advisor	AI recommends, human approves	Credit scoring, search ranking
AI as delegated agent	Human sets objectives, AI decides	Algorithmic trading, ad pricing
AI as autonomous actor	AI interprets/sets objectives	Frontier AI systems, AGI?

We are rapidly moving from Stage 2 → Stage 3.

Economics has thought a lot about Stages 1–2, but barely begun thinking about 3–4.

Aligned with Whom?

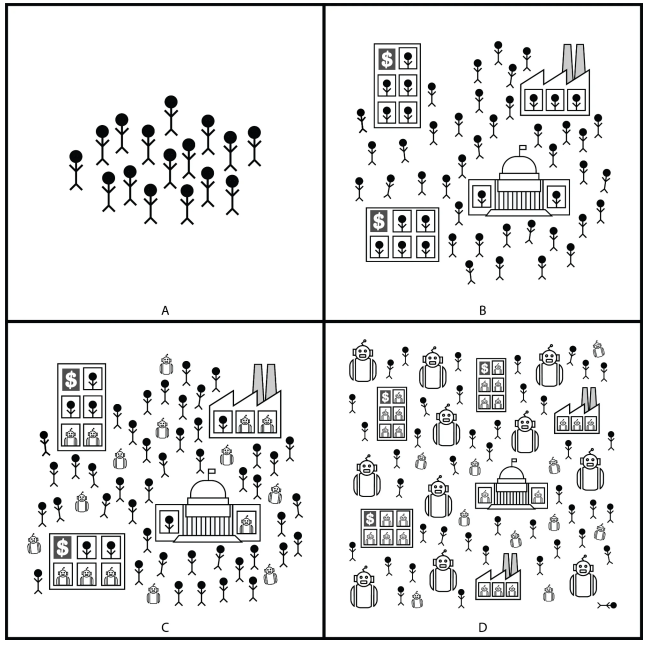
Growing AI agency raises the stakes on AI alignment

	Direct Alignment	**Social Alignment**
Question	Does the AI accomplish *the operator's* goals?	Does the AI serve *society's* goals?
Problem source	Technical: specification, communication, implementation	Governance: externalities, conflicting interests
Solution	Better engineering	Better institutions

Examples:

- Screening algorithm maximizes screening efficiency but encodes bias $\rightarrow$ *direct* ✓, *social* $\times$
- Recommendation engine maximizes engagement but polarizes $\rightarrow$ *direct* ✓, *social* $\times$

Source: Korinek & Balwit (2024), "Aligned with Whom? Direct and Social Goals for AI Systems," in Bullock et al. (eds.), *Oxford Handbook of AI Governance*, Oxford University Press


A
B
C
D