Strategic Content Creation: From Human-created Content to Generated Content

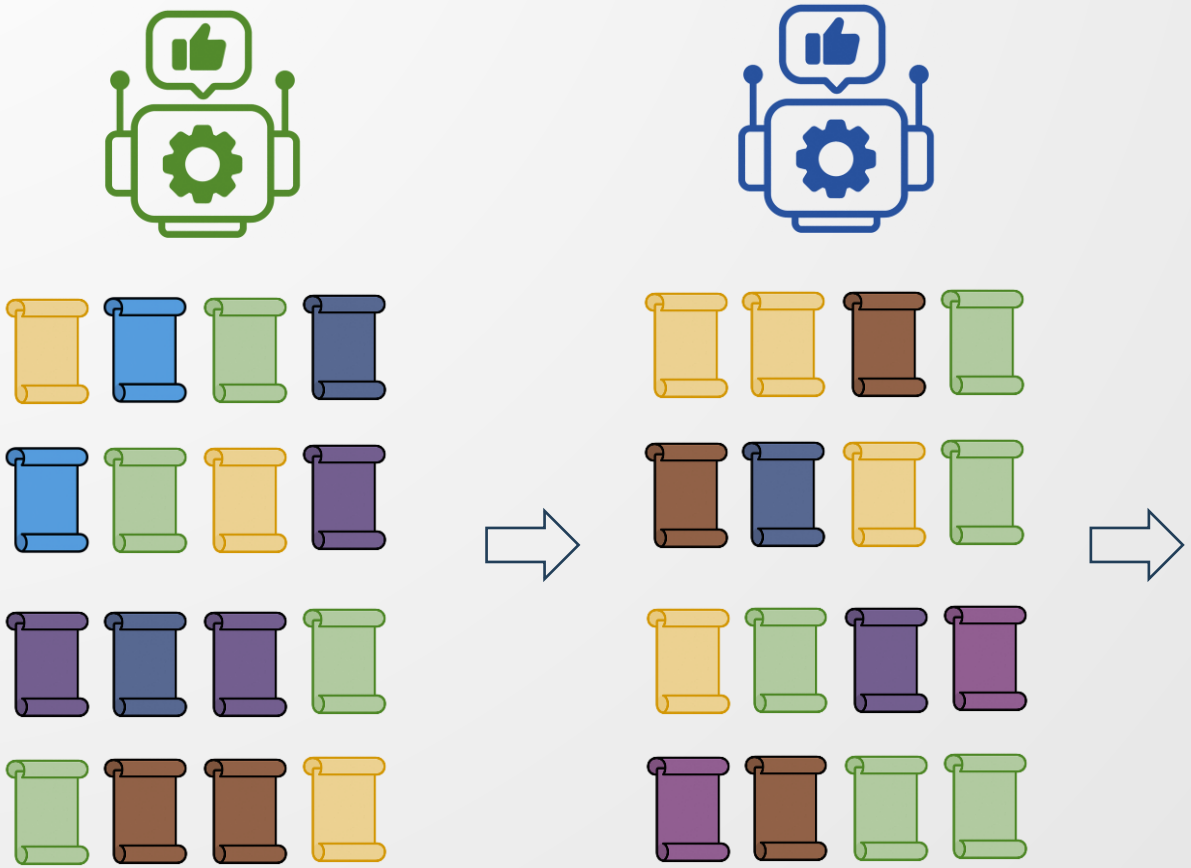
Omer Ben-Porat (Technion)

Robust IR @ SIGIR 2025

Robust IR

- *"We consider not only its average performance under normal conditions but, more importantly, its ability to maintain functionality across a variety of exceptional situations."*
- One such scenario: Strategic content creators
  - Even more since GenAI's rise
- Goals for this talk:
  - Establish motivation and raise challenges
  - W1: To Share or Not to Share? (In submission)
  - W2: Selective Response Strategies for GenAI (ICML 2025)

Strategic Behavior in RecSys

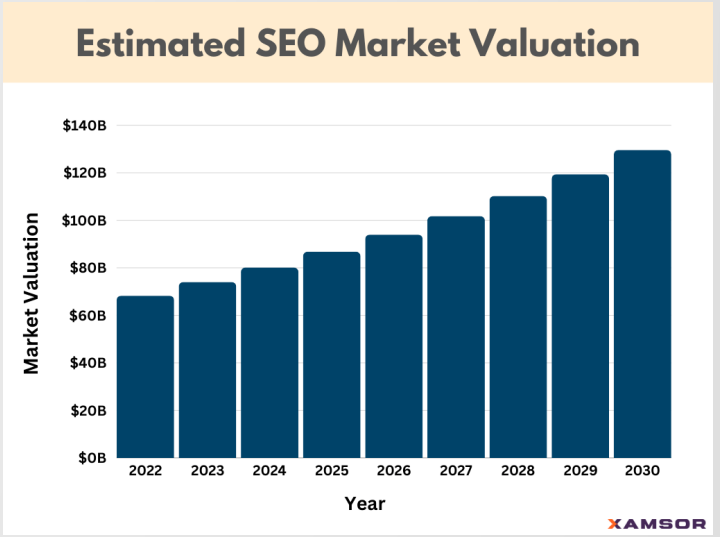


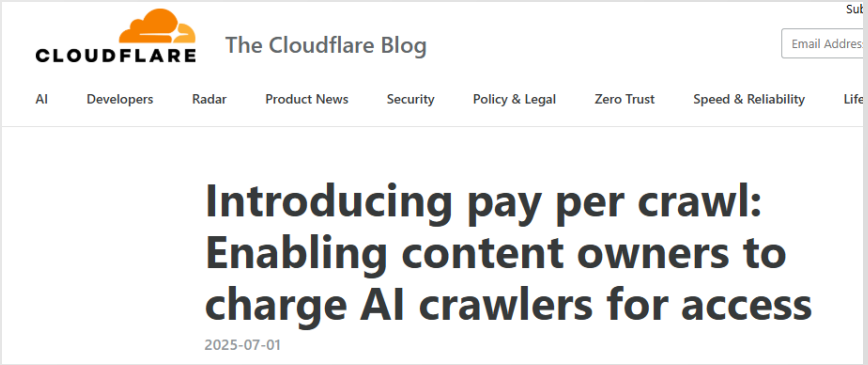
- SEO
- Promoting certain topics
- Promoting certain creators
- Changing the ranking function

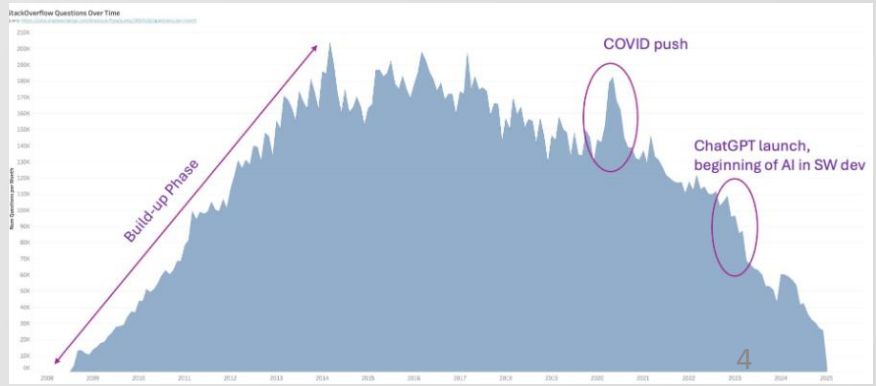
- E.g., Ben-Basat et al. 15', Ben-Porat & Tennenholtz '18, Ben-Porat et al. 19', Mladenov et al. '20, Jagadeesan et al. '22, Hron et al. '22, Calvano et al. '23, Castellini et al. '23, Ben-Porat & Torkan '23, Huttenlocher et al. '24

GenAI

- Generative engine optimization (GEO)
- Content can be created on demand
- Users are “locked” in the GenAI tool
  - Lower revenue for informative content
- Q&A forums get emptied
  - High-quality data is becoming scarcer

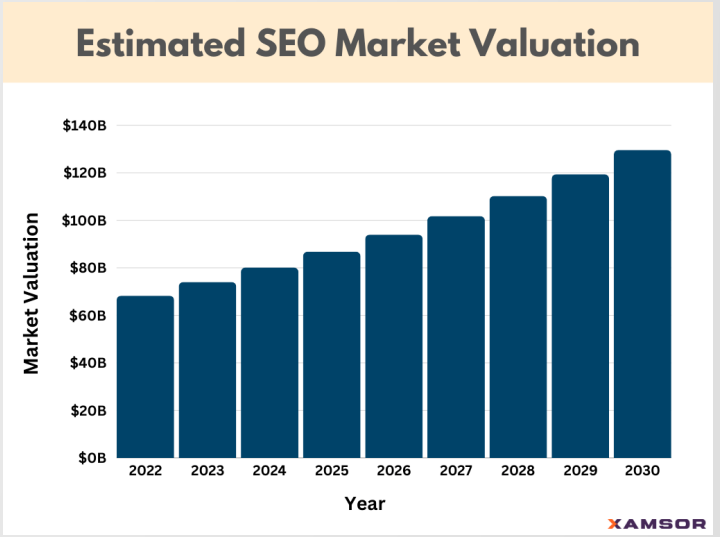


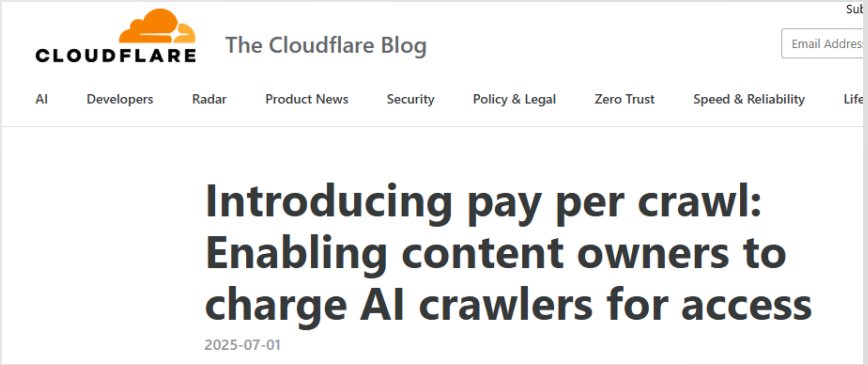


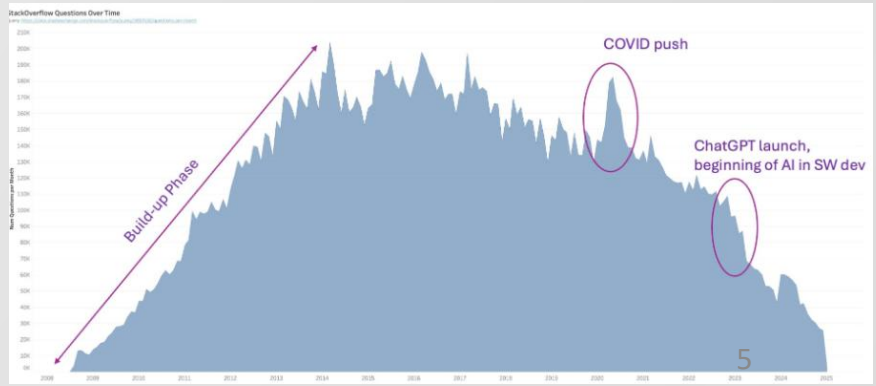


GenAI

- Generative engine optimization (GEO)
- Content can be created on demand
- Users are “locked” in the GenAI tool
  - Lower revenue for informative content
- Q&A forums get emptied
  - High-quality data is becoming scarcer

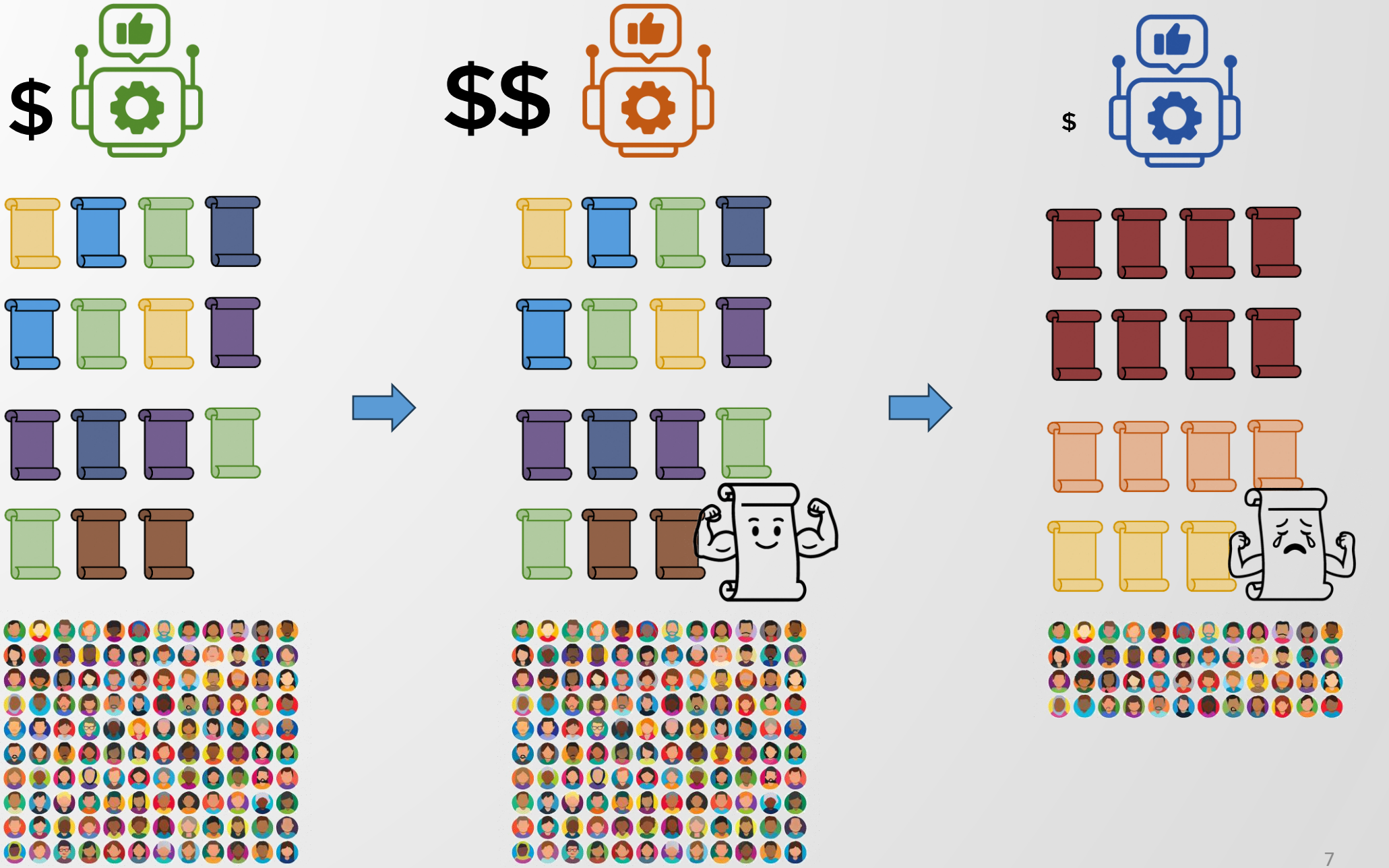






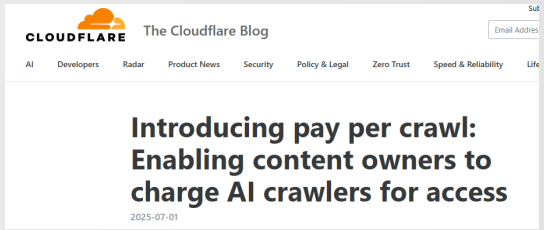
Strategic Content Creation in the Age of GenAI: To Share or Not to Share?

Gur Keinan and Omer Ben-Porat (in submission, link)



Idea: Distributing GenAI-Driven Revenue

- GenAI benefits with consumers and the platform, harms creators
- The platform CANNOT enforce sharing
  - Unauthorized usage, disputes,…
- Goals:
  1. Incentivize high quality content by creators
  2. (Human) Content sharing for better GenAI-based content
- Revenue allocation: Distribute the additional GenAI-driven revenue
  - Between creators who share content
- HOW?



Related Work

- Algorithmically mediated ecosystems
  - Incentivizing content quality
    - Hu et al. [2023], Immorlica et al. [2024], Yao et al. [2023a,b],
  - Ensuring fairness and diversity
    - Agarwal and Brown [2022], Dean and Morgenstern [2022], Mladenov et al. [2020], Yao et al. [2022a,b]
  - Aligning creator incentives with platform-level objectives
    - Boutilier et al. [2023], Zhu et al. [2023].
- Data valuation, data markets, copyright challenges
  - Baghcheband et al. [2024], Jia et al. [2019], Wang and Jia [2023], Acemoglu et al. [2022], Gans [2024], Pasquale and Sun [2024], Yang and Zhang [2024],
- Most related to ours: Yao et al. [2024a]
  - Similar traffic model, do not consider content sharing and revenue distribution

Model (1)

- n creators, strategy (x_i, s_i), representing (quality $\in [0, \infty)$, sharing level $\in [0,1]$)
- GenAI's quality $Q_{AI}(\boldsymbol{x}, \boldsymbol{s}) = \alpha \cdot \boldsymbol{x}^\top \boldsymbol{s}$
- Consumer traffic $T(\boldsymbol{x}) = \mu(\|\boldsymbol{x}\|_1)^\gamma$
- Probability of directing to content i (Tullock (1965))

$$\frac{x_i}{\|\boldsymbol{x}\|_1 + Q_{AI}(\boldsymbol{x}, \boldsymbol{s})}$$

- Probability of directing to GenAI

$$\frac{Q_{AI}(\boldsymbol{x}, \boldsymbol{s})}{\|\boldsymbol{x}\|_1 + Q_{AI}(\boldsymbol{x}, \boldsymbol{s})}$$

- Stackelberg: Platform commits to revenue distribution $\boldsymbol{f}$, then creators pick strategies

Model (2)

- Creator utility $U_i(\boldsymbol{x}, \boldsymbol{s}; \boldsymbol{f})$

$$T(\boldsymbol{x})\left(\frac{x_i}{\|\boldsymbol{x}\|_1 + Q_{AI}(\boldsymbol{x}, \boldsymbol{s})} + f_i(\boldsymbol{x}, \boldsymbol{s})\frac{Q_{AI}(\boldsymbol{x}, \boldsymbol{s})}{\|\boldsymbol{x}\|_1 + Q_{AI}(\boldsymbol{x}, \boldsymbol{s})}\right) - c_i(x_i)$$

- $T(\boldsymbol{x})$: Total traffic
- $\frac{x_i}{\|\boldsymbol{x}\|_1 + Q_{AI}(\boldsymbol{x}, \boldsymbol{s})}$: Prop. to i's content
- $f_i(\boldsymbol{x}, \boldsymbol{s})\frac{Q_{AI}(\boldsymbol{x}, \boldsymbol{s})}{\|\boldsymbol{x}\|_1 + Q_{AI}(\boldsymbol{x}, \boldsymbol{s})}$: GenAI's distributed Revenue
- $c_i(x_i)$: Content production costs

Model (2)

- Creator utility $U_i(\boldsymbol{x}, \boldsymbol{s}; \boldsymbol{f})$

$$T(\boldsymbol{x})\left(\frac{x_i}{\|\boldsymbol{x}\|_1 + Q_{AI}(\boldsymbol{x}, \boldsymbol{s})} + f_i(\boldsymbol{x}, \boldsymbol{s})\frac{Q_{AI}(\boldsymbol{x}, \boldsymbol{s})}{\|\boldsymbol{x}\|_1 + Q_{AI}(\boldsymbol{x}, \boldsymbol{s})}\right) - c_i(x_i)$$

- Platform's revenue $U_P(\boldsymbol{x}, \boldsymbol{s}; \boldsymbol{f})$

$$T(\boldsymbol{x})\frac{Q_{AI}(\boldsymbol{x}, \boldsymbol{s})}{\|\boldsymbol{x}\|_1 + Q_{AI}(\boldsymbol{x}, \boldsymbol{s})}\left(1 - \sum_{i=1}^{N} f_i(\boldsymbol{x}, \boldsymbol{s})\right)$$

Traffic

Directly to GenAI

Undistributed Revenue

Model (2)

- Creator utility $U_i(\boldsymbol{x}, \boldsymbol{s}; \boldsymbol{f})$

$$T(\boldsymbol{x})\left(\frac{x_i}{\|\boldsymbol{x}\|_1 + Q_{AI}(\boldsymbol{x}, \boldsymbol{s})} + f_i(\boldsymbol{x}, \boldsymbol{s})\frac{Q_{AI}(\boldsymbol{x}, \boldsymbol{s})}{\|\boldsymbol{x}\|_1 + Q_{AI}(\boldsymbol{x}, \boldsymbol{s})}\right) - c_i(x_i)$$

- Platform's revenue $U_P(\boldsymbol{x}, \boldsymbol{s}; \boldsymbol{f})$

$$T(\boldsymbol{x})\frac{Q_{AI}(\boldsymbol{x}, \boldsymbol{s})}{\|\boldsymbol{x}\|_1 + Q_{AI}(\boldsymbol{x}, \boldsymbol{s})}\left(1 - \sum_{i=1}^{N} f_i(\boldsymbol{x}, \boldsymbol{s})\right)$$

- GenAI's revenue distribution, parametrized by ρ

$$f_{i,\rho}(\boldsymbol{x}, \boldsymbol{s}) = \begin{cases} \rho\dfrac{x_i s_i}{\boldsymbol{x}^\top \boldsymbol{s}} & \text{if } \boldsymbol{x}^\top \boldsymbol{s} > 0 \\ 0 & ow \end{cases}$$

No Revenue Distribution?

- What happens if the platform sets $\rho = 0$?
  - The platform keeps all the GenAI-driven revenue
- Creators only get their share:

$$U_i(\boldsymbol{x}, \boldsymbol{s}; \rho = 0) = T(\boldsymbol{x}) \frac{x_i}{\|\boldsymbol{x}\|_1 + \cancel{Q_{AI}(\boldsymbol{x}, \boldsymbol{s})}} - c_i(x_i)$$

- Thus, creators will decide not to share!
  - -> no GenAI-driven revenue for the platform

Depends on the creators!

- We need to analyze creator incentives in the game $G(\rho)$

Nash Equilibrium

- We say a profile $\boldsymbol{x} = (x_1, \dots, x_i, \dots, x_n), \boldsymbol{s} = (s_1, \dots, s_i, \dots, s_n)$ is a pure Nash equilibrium (PNE) in $G(\rho)$ if for every creator i

$$U_i\big((x_i, \boldsymbol{x}_{-i}), (s_i, \boldsymbol{s}_{-i}); \rho\big) \geq \sup_{x_i', s_i'} U_i\big((x_i', \boldsymbol{x}_{-i}), (s_i', \boldsymbol{s}_{-i}); \rho\big)$$

$\boldsymbol{x}_{-i} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$

$\boldsymbol{s}_{-i} = (s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n)$

Nash Equilibrium

- We say a profile $\boldsymbol{x} = (x_1, \ldots, x_i, \ldots, x_n), \boldsymbol{s} = (s_1, \ldots, s_i, \ldots, s_n)$ is a pure Nash equilibrium (PNE) in $G(\rho)$ if for every creator i

$$U_i\big((x_i, \boldsymbol{x}_{-i}), (s_i, \boldsymbol{s}_{-i}); \rho\big) \geq \sup_{x_i', s_i'} U_i\big((x_i', \boldsymbol{x}_{-i}), (s_i', \boldsymbol{s}_{-i}); \rho\big)$$

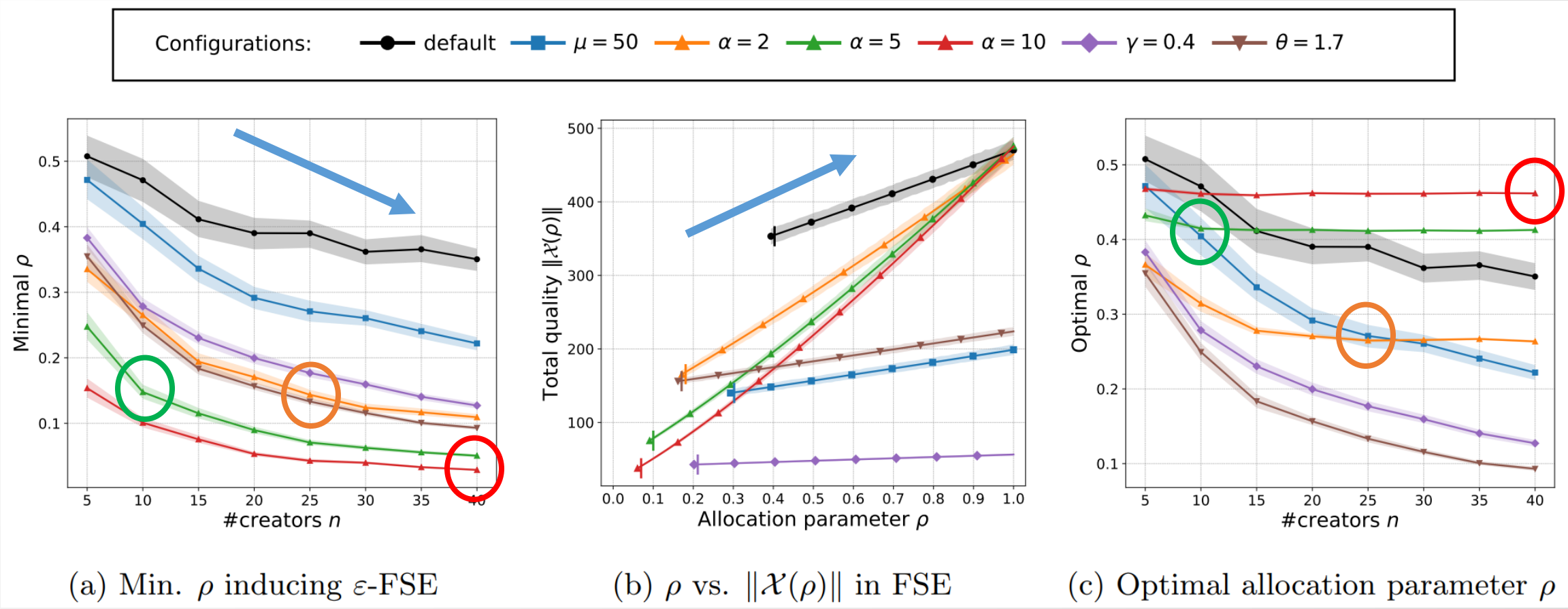
- In words, (x_i, s_i) is a *best response* to $(\boldsymbol{x}_{-i}, \boldsymbol{s}_{-i})$ in $G(\rho)$
- If $\boldsymbol{x} = (x_1, \ldots, x_n), \boldsymbol{s} = \mathbf{1}$ is a PNE, we call it a full sharing PNE (FSE)
- Highly desirable!
  - Maximizes $Q_{AI}(\boldsymbol{x}, \boldsymbol{s}) = \alpha \cdot \boldsymbol{x}^\top \boldsymbol{s}$ given $\boldsymbol{x}$
  - Reduces disputes from creators over unconsented content usage
- **Theorem**: If ρ is large "enough", there exists a unique FSE
  - 1%-50% in "standard" settings, decreasing with n
  - Prisoner's dilemma-like phenomenon

Revenue Optimization

- How to pick ρ?

$$\max_{\rho \in [0,1]} U_P(\boldsymbol{x}, \mathbf{1}; \rho)$$
$$\text{subject to } (\boldsymbol{x}, \mathbf{1}) \text{ is the unique FSE in } G(\rho)$$

- Bi-level optimization are NP-hard
- We develop an efficient approximately optimal algorithm
  - Smoothness analysis of fixed points, Implicit Function Theorem

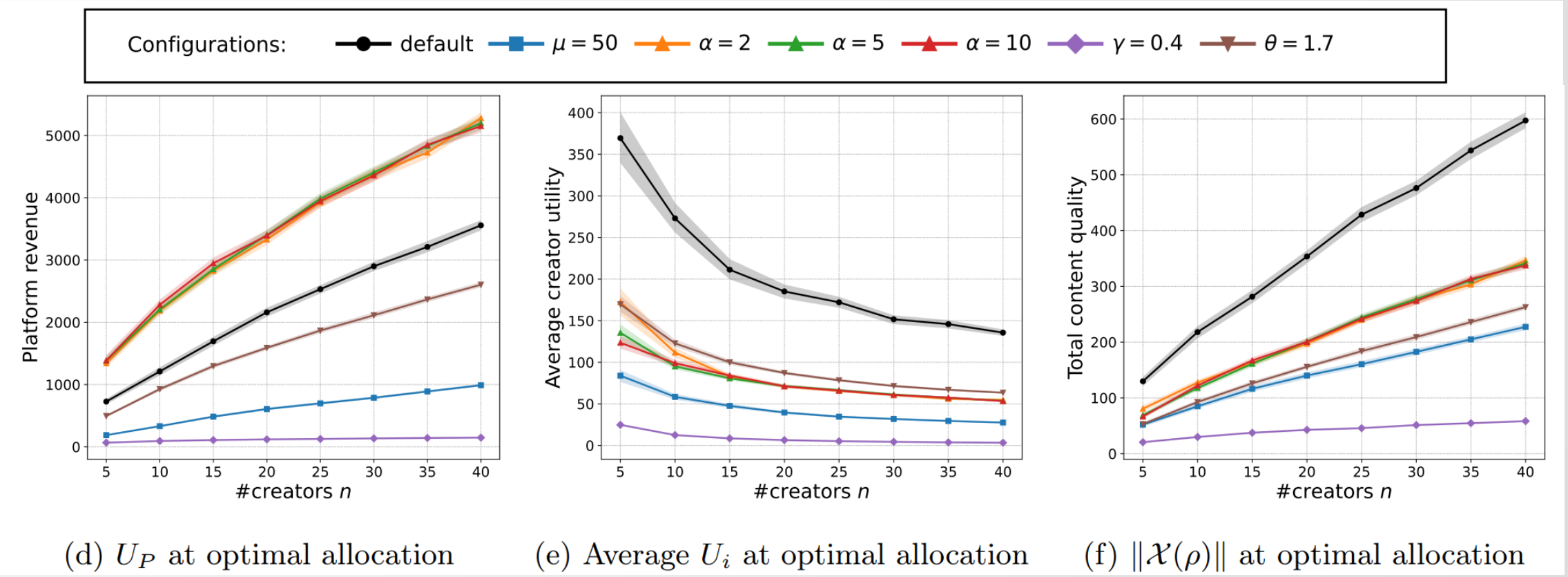


(a) Min. ρ inducing ε-FSE

(b) ρ vs. $\|\mathcal{X}(\rho)\|$ in FSE

(c) Optimal allocation parameter ρ

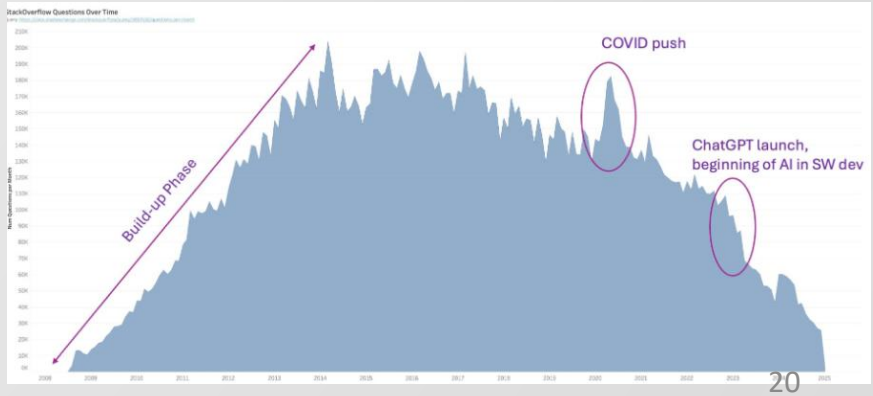
1. Minimal values of ρ reduce with #creators
2. Total quality increases with ρ
3. Minimal is not (necessarily) optimal



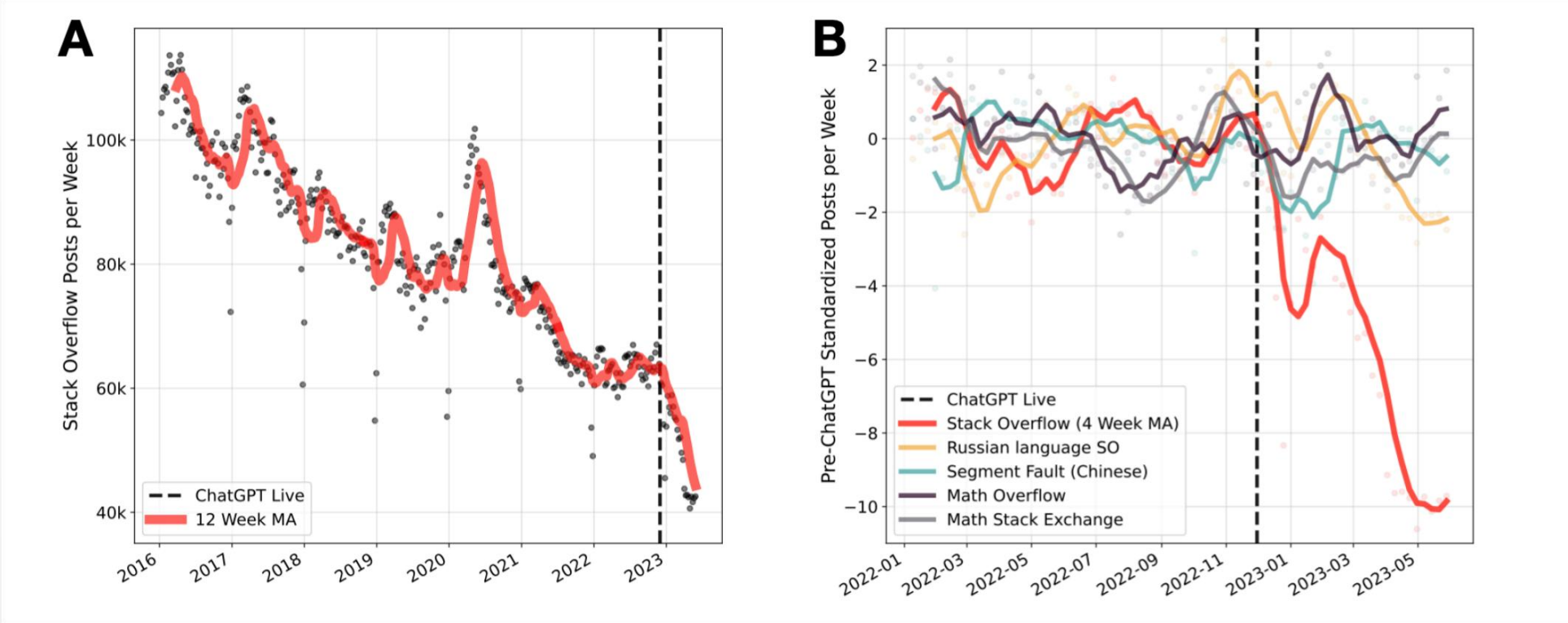
(d) U_P at optimal allocation (e) Average U_i at optimal allocation (f) $\|\mathcal{X}(\rho)\|$ at optimal allocation

Selective Response Strategies for GenAI

Boaz Taitler and Omer Ben-Porat (ICML 25', link)

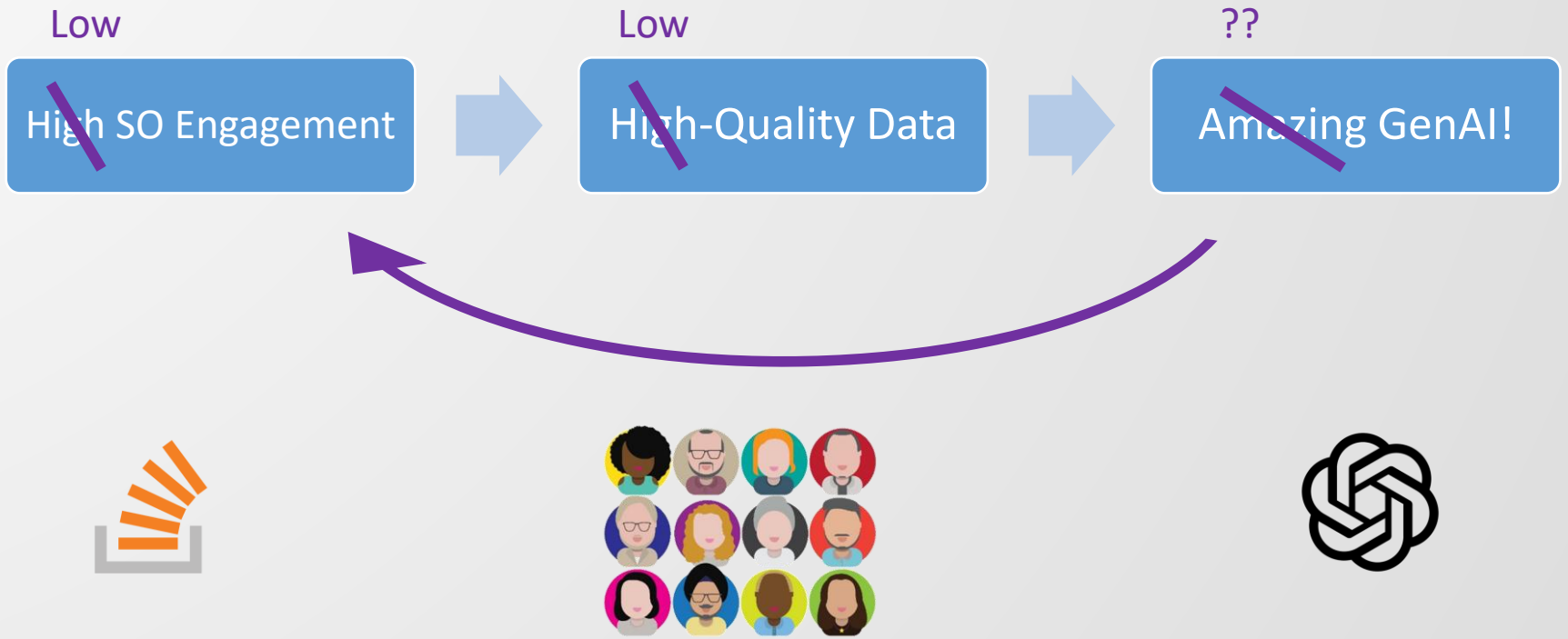


Motivation



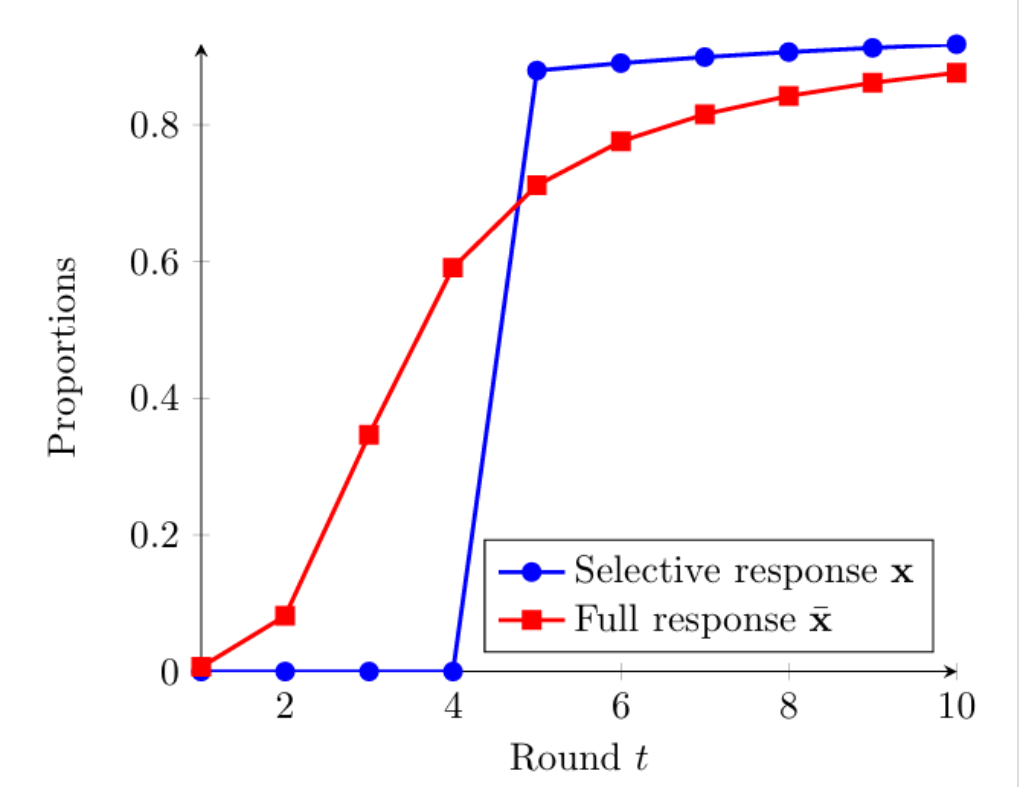
del Rio-Chanona, M., Laurentsyeva, N., & Wachs, J. (PNAS Nexus 2024). Are Large Language Models a Threat to Digital Public Goods? Evidence from Activity on Stack Overflow.

Content Dynamics and GenAI



Idea: Selective Response

- New topic, technology, entity (no “training” data)
- GenAI doesn’t “know too much”
- GenAI can respond “selectively”
  - (abstract)
- Users generate data whenever they don’t use GenAI
  - Externally (Forums) or internally (Base44 is wrong->you fix)

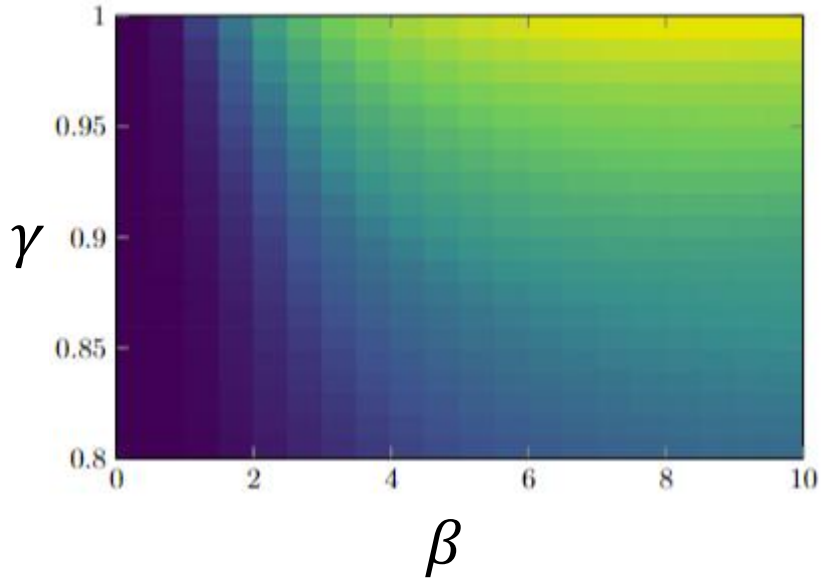


Related Work

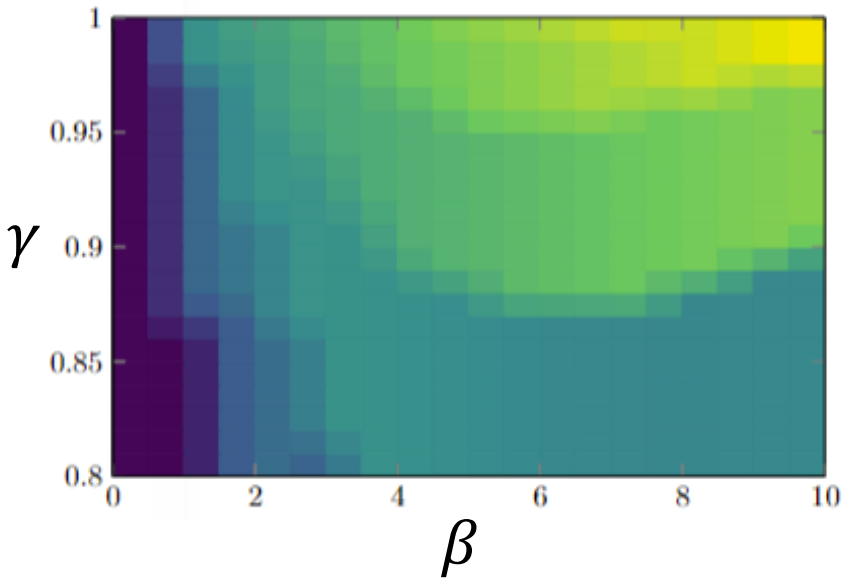
- Foundation models X game theory
  - Raghavan [2024], Laufer et al. [2024], Conitzer et al. [2024], Dean et al. [2024], Dütting et al. [2024], Taitler and Ben-Porat [2025]
- Information design
  - Strategic disclosure of information (Bergemann and Morris [2019], Bergemann et al. [2015])
  - Strategic communication (Crawford and Sobel [1982], Milgrom [1981], Babichenko et al. [2023], Lu et al. [2023])
  - Cheap talk (Lo et al. [2023], Crandall et al. [2018])
- Competition between platforms
  - Rietveld and Schilling [2021], Karle et al. [2020], Bergemann and Bonatti [2024], Tullock [1980].
  - Competition in marketplaces (Jagadeesan et al. [2023], Feldman et al. [2013])
  - Competing Bandits (Aridor el al. [2025])

Price-of Analysis

Revenue



Welfare



+Approximately optimal algorithms!

Conclusions

- Content sharing with GenAI: A different era
- Game theory: Reasoning about the consequences of a given recommendation policy
- Today: Theory, TRL 1
- Many limitations
- Happy to elaborate and collaborate ☺