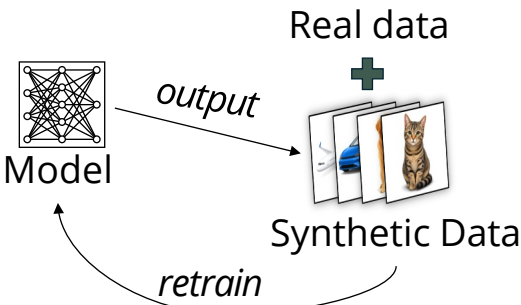


Core Question: *Stability · Diversity · Alignment · Competition*

How do **feedback loops** among **models, data, rewards, users, and markets** shape the **long-term behavior** of generative AI ecosystems?

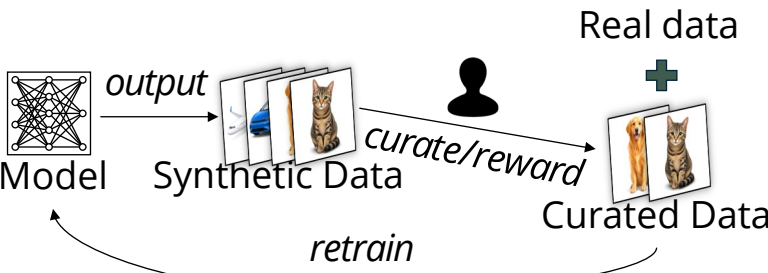
Existing Foundations

Self-Consuming Model



- A model retrains *on its own synthetic outputs*.
- Repeated self-consumption may cause **drift, diversity loss, and collapse**.

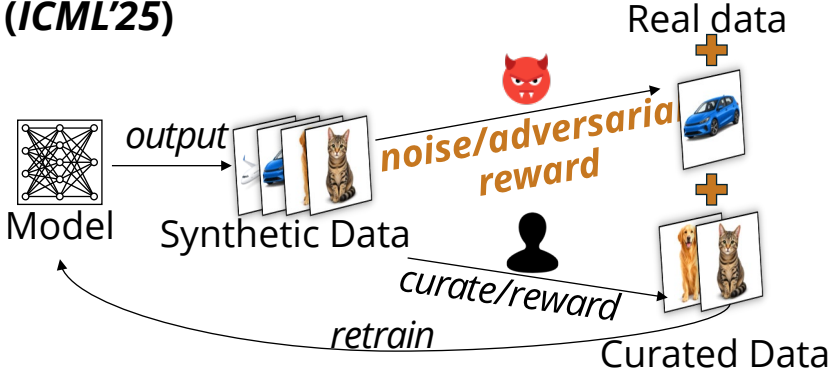
Self-Consuming + Reward/Curation



- Users (reward) *select synthetic samples* for retraining.
- Curation steers the model toward **preferred or high-reward** attributes.

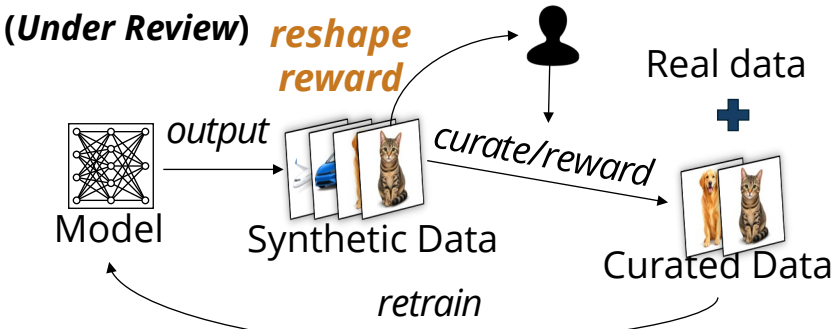
My Work

Self-Consuming + Adversarial Curation (ICML'25)



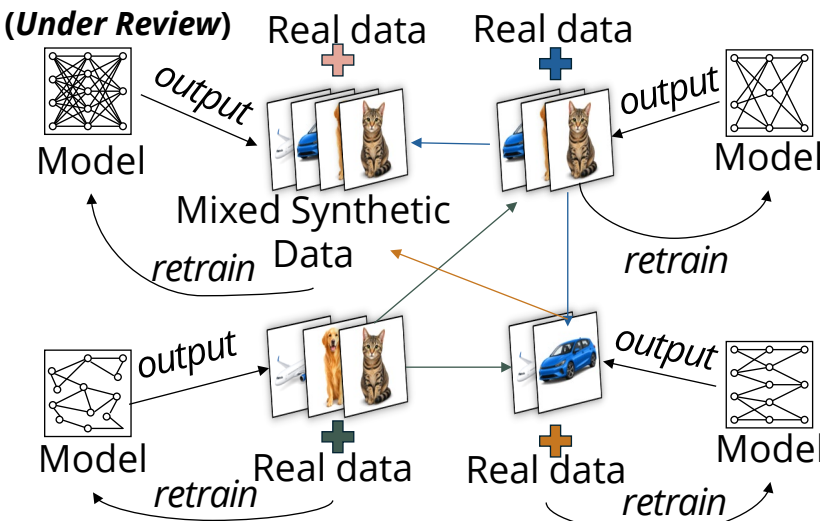
- Self-consuming retraining is *robust to mild noise*, but strategic curation can **misalign future models**.
- We design budget-limited attacks that steer models away from true user preferences.

Self-Consuming + Co-Evolving Reward (Under Review)



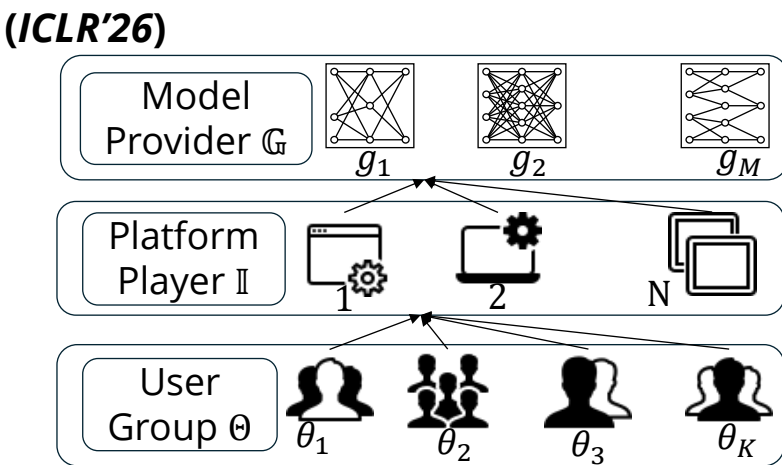
- Model outputs reshape user preferences, creating **preference lock-in**.
- We design reference-data interventions to *preserve desired attributes*.

Networked Self-Consuming models (Under Review)



- **Network topology** and feedback cycles determine model *stability and diversity*.

Market Games for Generative Models (ICLR'26)



- **Competition** shapes *differentiation, welfare, and market diversity*.