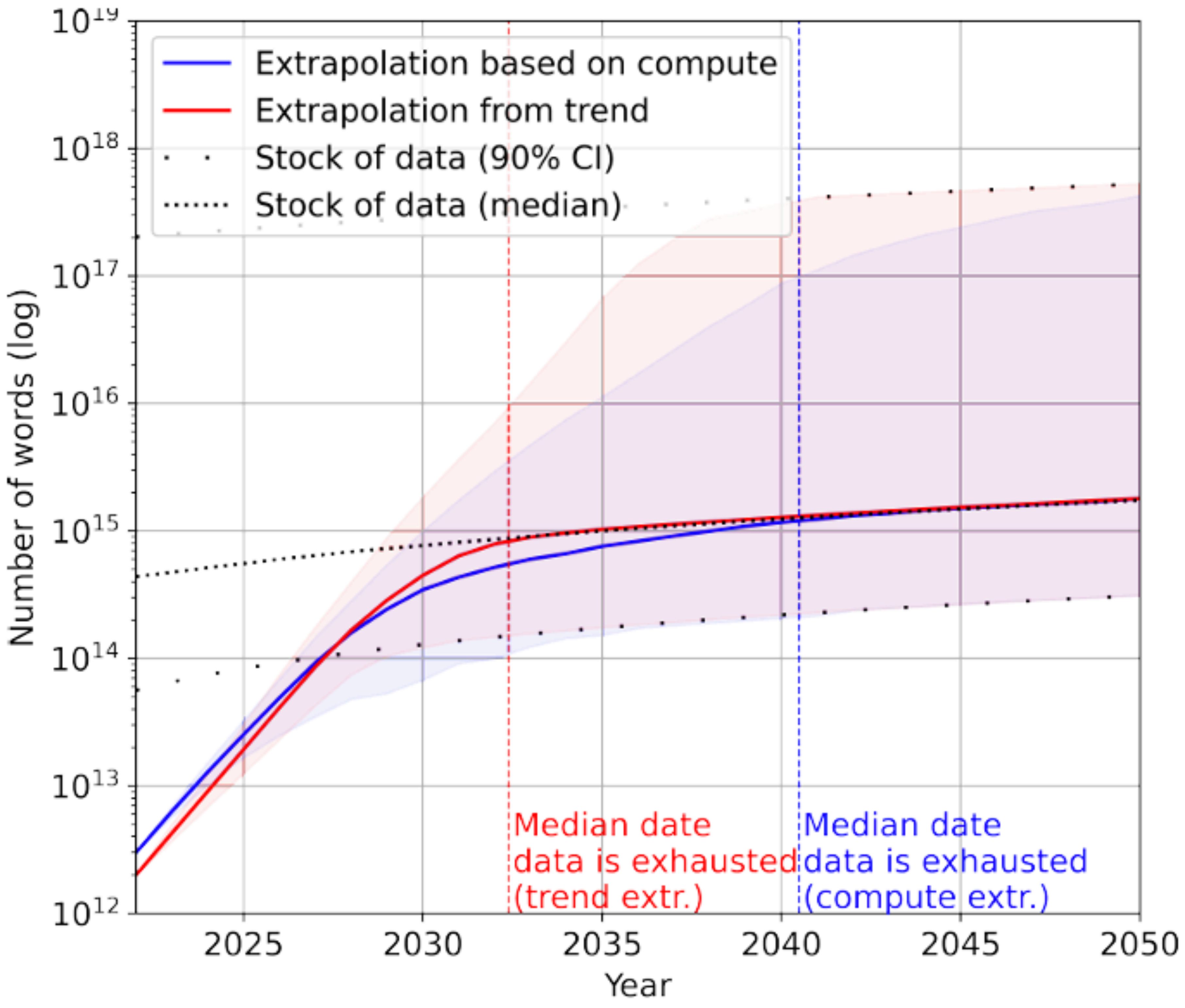
Trustworthy Federated Learning with Untrusted Participants

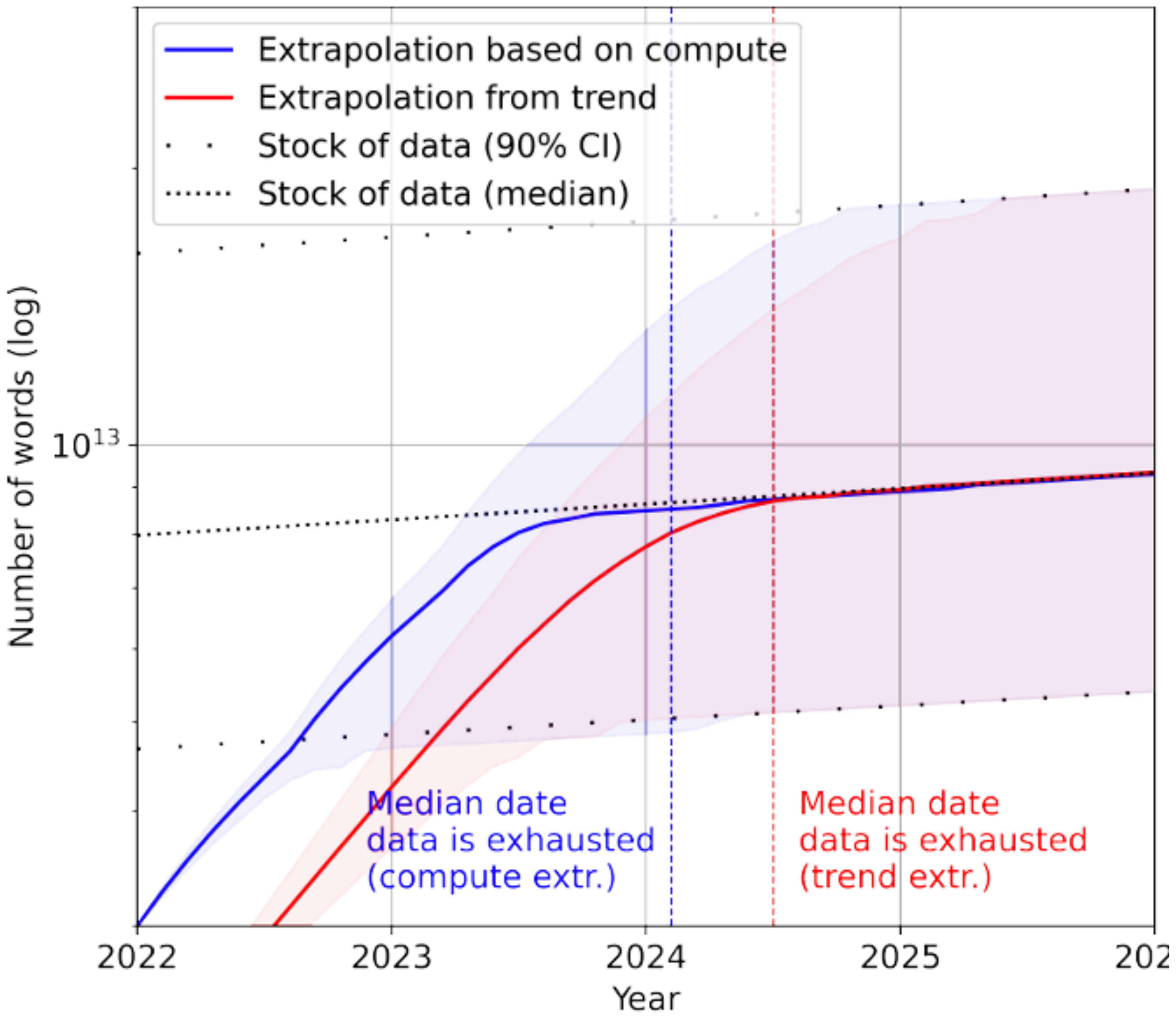
Youssef Allouah, Rachid Guerraoui, John Stephan

June 2025

Not Enough (Nice) Data



(a) Projections for low-quality language data



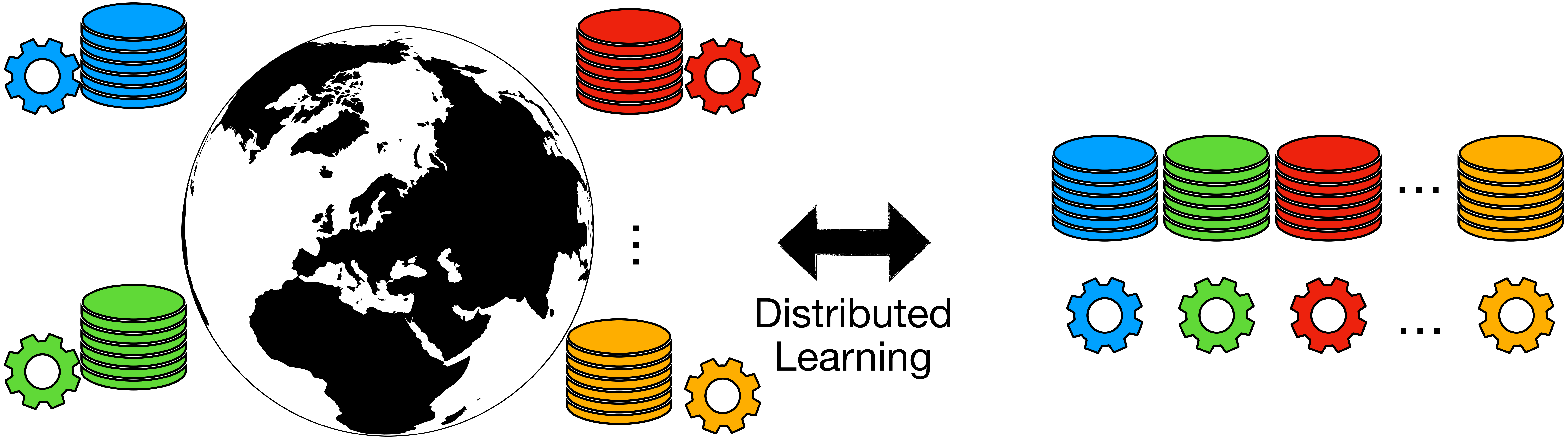
(b) Projections for high-quality language data

Villalobos et al. 2022

Not Enough Compute (Eventually)

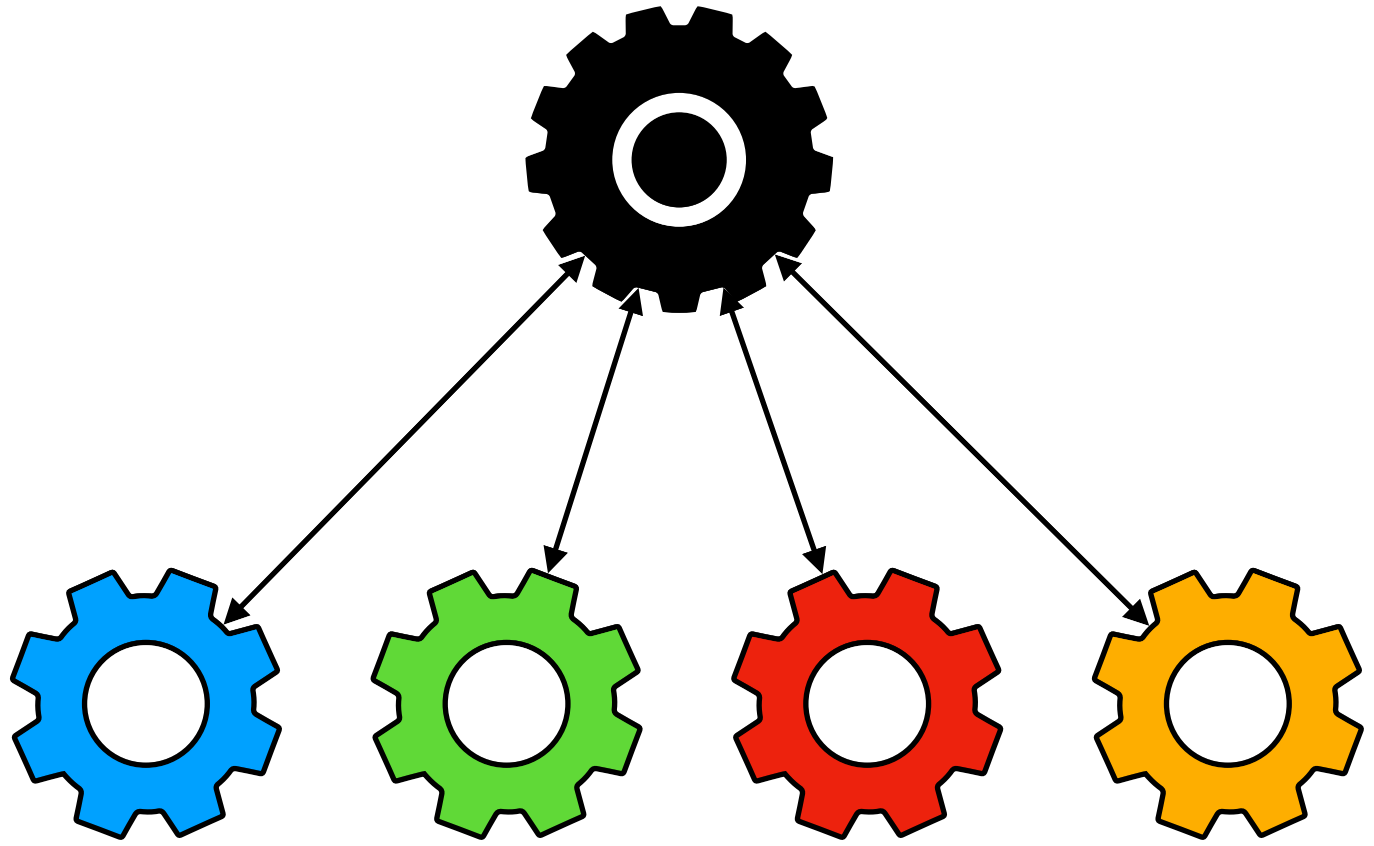
Distributed Learning Promise

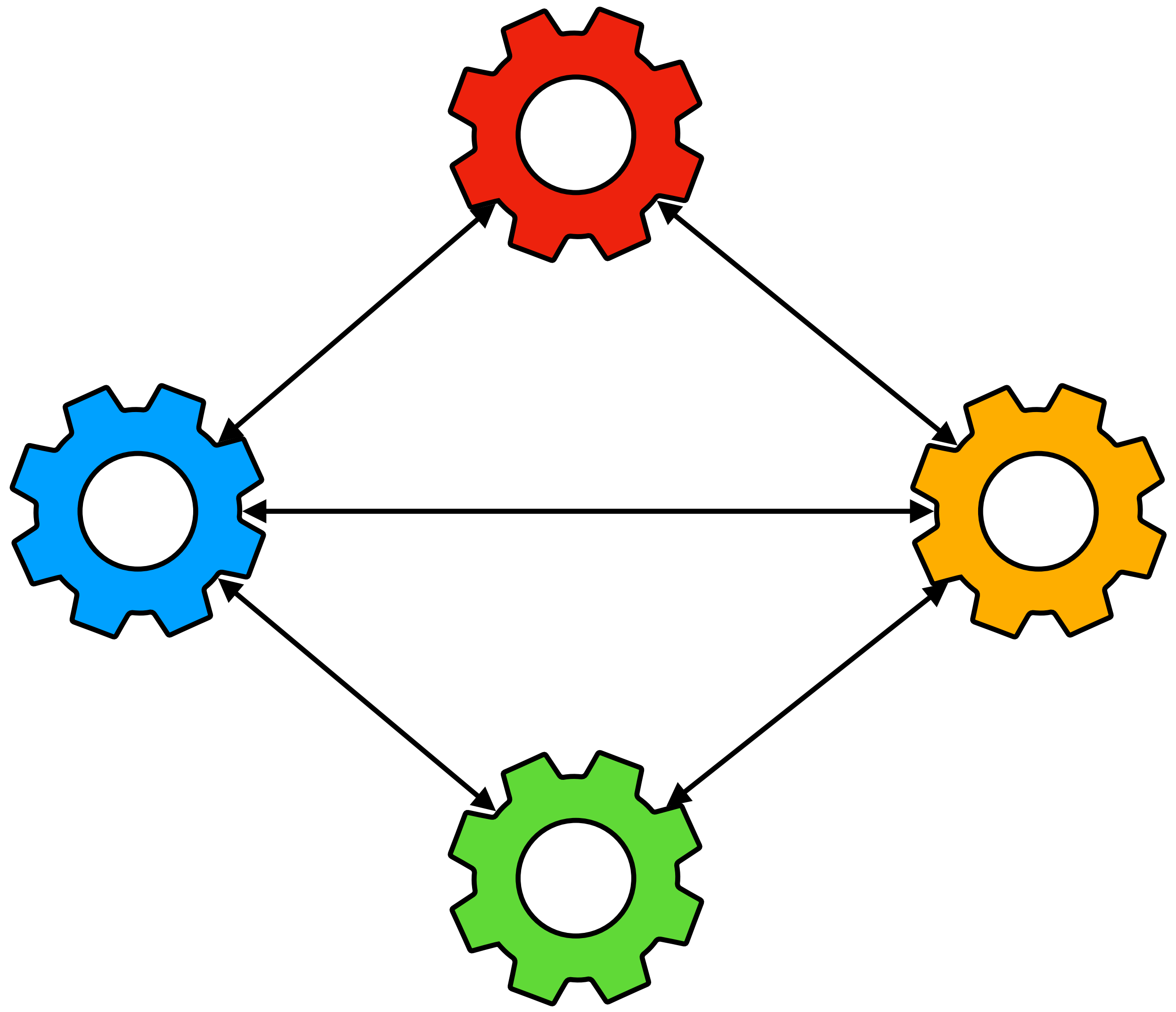
- Modern Machine Learning: needs large models, massive datasets
- Distributed Learning: keep data local, offload compute



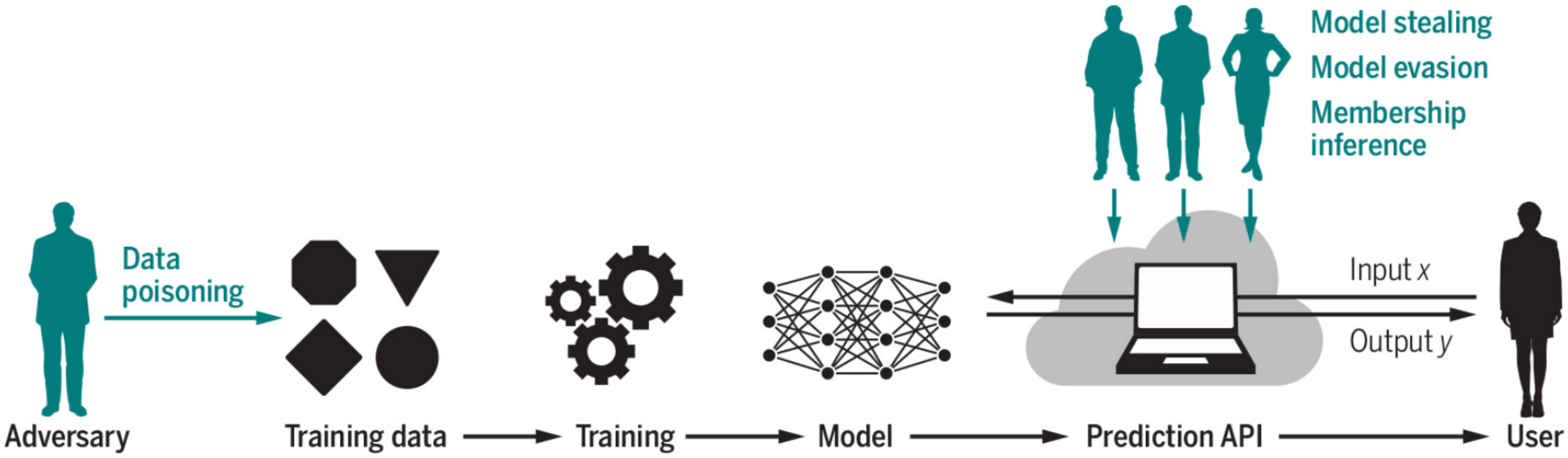
Distributed Learning Promise

- System examples: federated, decentralized/peer-to-peer





Threats in Machine Learning



Eshete, 2021

Threats in Distributed Learning

- Challenges: model poisoning, communication/network faults, hardware failures, data privacy, …

Threats in Distributed Learning

- Challenges: model poisoning, communication/network faults, hardware failures, data privacy, …
- **Corruption** adversary (Byzantine): controls fraction of workers, full knowledge, computationally unbounded — aims to disable learning
- **Privacy** adversary: observes all communications and models — aims to infer membership/extract data
- Many others: fairness, data ownership, …

State of the Art

- Robustness and Privacy induce a **coupled hardness**: requires specific solutions, motivates weaker threat models

  [AGGPS, ICML ’23] “On the Privacy-Robustness-Utility Trilemma in Distributed Learning”

*almost

State of the Art

- Robustness and Privacy induce a **coupled hardness**: requires specific solutions, motivates weaker threat models [AGGPS, ICML ‘23]
- Spectrum of weaker threat models:
  - **Trusted Server:** central DP privacy-utility trade-off
  - **Trusted Shuffler:** central* DP trade-off
  - **Comput. boundedness:** central* DP trade-off, requires cryptography
  - **Shared Randomness:** near-central DP trade-off, no* cryptography

*almost

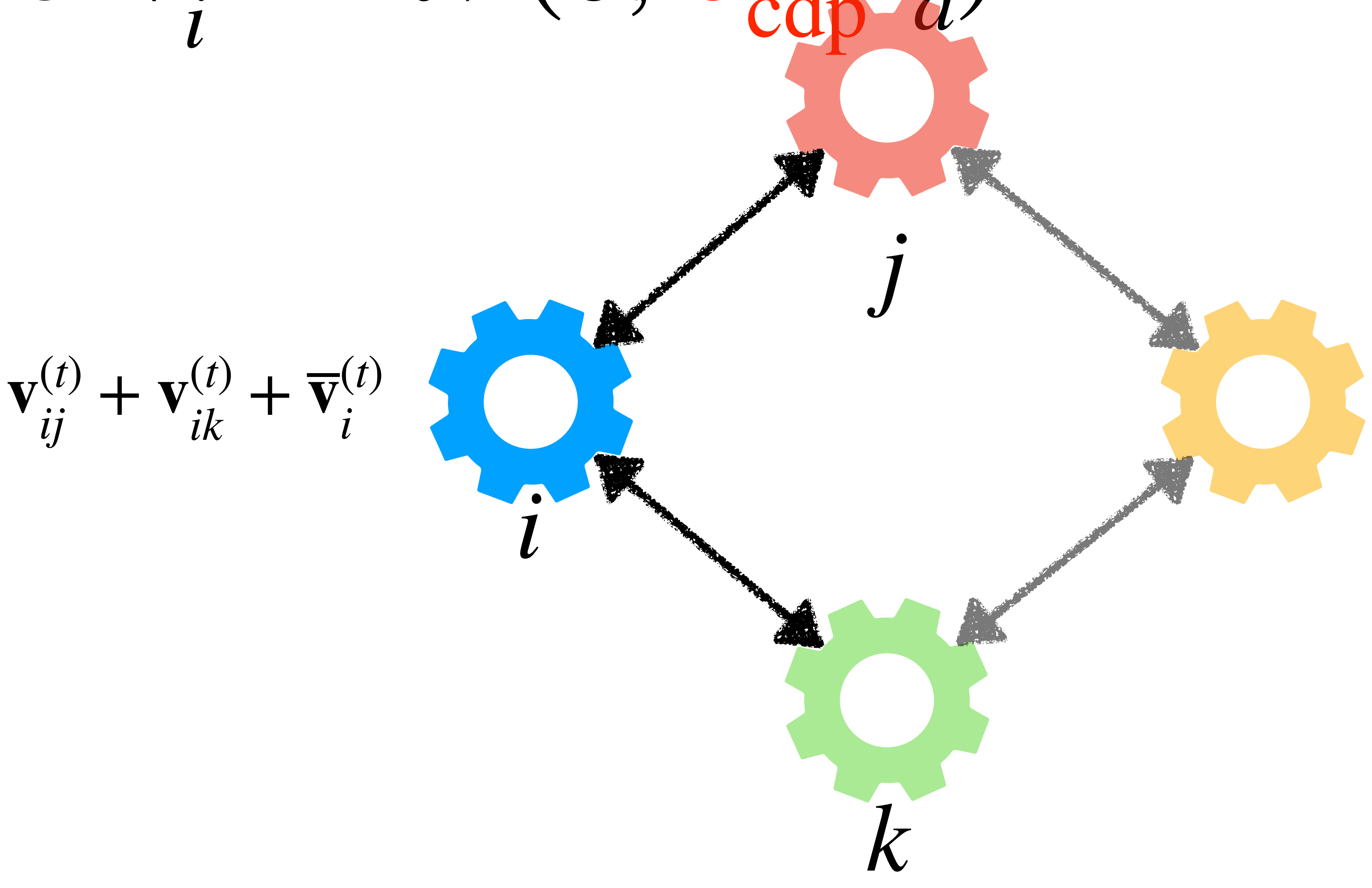
Algorithmic Ideas

Correlated Noise

- Pairwise canceling shares: $\mathbf{v}_{ij}^{(t)} = -\mathbf{v}_{ji}^{(t)} \sim \mathcal{N}(0,\ \sigma_{\text{cor}}^2 \mathbf{I}_d), \forall$neighbors i, j

Bonawitz et al., 2017

- Uncorrelated share: $\overline{\mathbf{v}}_i^{(t)} \sim \mathcal{N}(0,\ \sigma_{\text{cdp}}^2 \mathbf{I}_d)$



Algorithmic Ideas

High-dimensional robust aggregation

- CAF aggregation efficiently looks at all dimensions at once; adapts spectral filtering ideas from TCS community [DKKLMS, FOCS ’16]
- High-dimensional robustness is crucial: correlated noise amplifies the vulnerability to malicious participants
- Variance-reduction across iterations: we use local-client momentum to improve robustness [KHJ, ICML ’21] [FGGPS, ICML ’22]

Open Questions

1. More **utility**: what is the best privacy-utility trade-off using shared randomness only?

2. More **efficiency**: What is the best computational and communication complexity we can achieve for the same privacy-utility trade-off?

X/Twitter: @ys_alh

LinkedIn: Youssef Allouah

youssef.allouah@epfl.ch