Having Your Privacy Cake and Eating it Too: Platform-supported Auditing of Social Media Algorithms for Public Interest

Basileal Imana¹, Aleksandra Korolova² and John Heidemann¹
¹ USC/Information Sciences Institute ² Princeton University

Motivation

- Increasing desire for scrutiny of social-media algorithms
- Digital Services Act (EU) and PATA (USA)
- Concerns regarding algorithmic personalization
- Bias/discrimination, political polarization, …
- Tension in value for different stakeholders
  - Users: personalized content => positive experience
  - Businesses: increased engagement => maximize profit
  - Society: fairness => not promote bias/polarization

Our Contributions

1) Discuss the needs for a new type of algorithmic auditing
2) Suggest query access to relevance estimators is the key for increasing transparency
3) Show such access need not risk user privacy nor business interests
4) Quantify number of samples needed for auditing with privacy

Motivation

- Increasing desire for scrutiny of social-media algorithms
- Digital Services Act (EU) and PATA (USA)
- Concerns regarding algorithmic personalization
- Bias/discrimination, political polarization, …
- Tension in value for different stakeholders
  - Users: personalized content => positive experience
  - Businesses: increased engagement => maximize profit
  - Society: fairness => not promote bias/polarization

Conclusion

- A path exists from proposed legislation to realizable auditing system
- Platforms can enable platform-supported auditing without compromising privacy
- Next steps:
  - Collaborate with an ad platform to implement the framework
  - Paper pre-print: https://arxiv.org/abs/2207.08773
  - To appear in CSCW ’23

This work was funded in part by a grant from The Rose Foundation and NSF awards 1916153, 1956435, and 1943584.