Supplemental Bibliography for “Learning to Live with Privacy-Preserving Analytics”


This bibliography includes additional references for:

- Privacy-preserving analytics techniques, as a subset of privacy-enhancing technologies [22]: blind signatures [8], homomorphic encryption [21], multi-party computation [52], k-anonymity [49], differential privacy [12,14,15] (including differentially private machine learning [1] and local differential privacy [29]), and federated learning [34] [33,42]. Re-identification [9], and the Census Bureau’s re-identification attack [2]. As compared to traditional statistical disclosure limitation [19].
- Applications mentioned: Census Bureau’s OnTheMap tool [31], Google’s initial FLoC proposal [6,23], Facebook’s URLs dataset [41], HIPAA [16,17], the Facebook advertising stack [35], various internal SQL systems at Google, Uber, Oracle, SAP, and other tech companies [11,24], local telemetry in Apple devices [25], Google’s GBoard keyboard [34], and a LinkedIn research dataset [44]. Also potential applications for platform auditing [4].
- Opposition to privacy plans around the census [7,37]: from scholars [28,30,45], politicians [53], and civil rights groups [32,40]. Also, pushback from advertisers on Google’s third party cookie replacement [38].
- Utility-privacy trade-offs [3,13,20,36,50], including concerns about impacts on vulnerable groups [5,27,45,48] or hard-to-count groups [18,51], industry norms [14], and stakeholders’ preferences [10].
- Studies by Sarathy et al. [46] and others [14,26] describe practitioners’ struggles with understanding and successfully applying PPA tools.
- From our study [48], estimates of the amount of census-guided federal funding [43] and the Title I program specifically [47].
- Other notable discussions of PPAs and their implications for social science: Oberski & Kreuter [39], Hotz et al. [28], and Abowd et al. [4].


