**More Differential Privacy!**

<https://www.youtube.com/watch?v=_g83jr1LChA>

**Differential Privacy and the People’s Data**

Population as a Whole vs Needle in a Haystack

* DP preserves the privacy of individuals while allowing for statistical information to be extracted from a data set.
* DP is “entirely the wrong tool for the needle in the haystack” [4:09].
  + DP protects all of the outliers of the data set and “hides the needle, just allowing the shape of the haystack to come through”.

Statistics “feel” private

* A quantity computed from a sample tells us about the population as a whole.
* Two datasets, collected independently from the same population and with correct methodology, should in essence tell the same things.
* Sense of privacy derived from this fact.
  + “no one knows I was in the sample”
  + “I can claim I opted out”
  + “it’s not about me”
* On the right track but needs some help. DP provides this help.

DP extends “Statistical” Privacy to all computations

Privacy-Preserving Data Analysis

<https://www.youtube.com/watch?v=O97jAe_pXdo&t=39s>

<https://www.youtube.com/watch?v=zNm6efQI0h8>

<https://www.microsoft.com/en-us/research/publication/platform-for-differential-privacy/>

Differential privacy works great for large data sets, but what about data sets that contain less than 100 data entries?