

The
Alan Turing
Institute

Quantifying Uncertainty and Privacy Preservation in Synthetic Data

Martin O'Reilly



The problem



Health



Finance



Census



Location



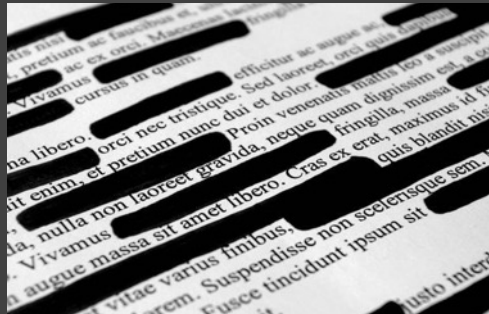
Control data access



Make data safer



Summarise



Anonymise



Sample



Perturb



Remove



Synthesise



?

?

?

?

?

?



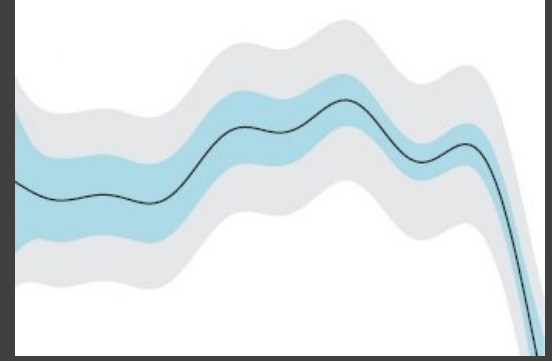
What are we doing?



Tools



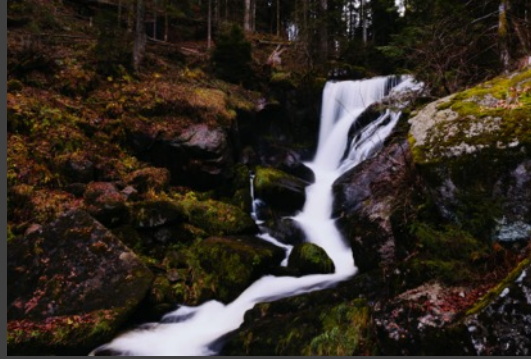
Privacy vs utility



Uncertainty



Data structure



Data source



Analysis question

How can we help each other?



Sensitive datasets



Analysis questions



Synthesis techniques



Martin
O'Reilly



Alison
Heppenstall



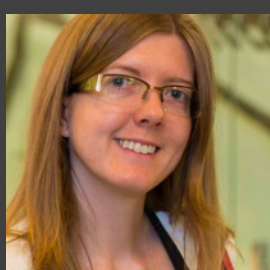
Nik Lomax



Sebastian
Vollmer



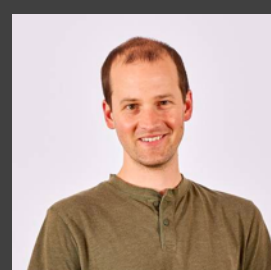
Oliver
Strickson



Louise
Bowler



Kasra
Hosseini



Nick
Malleson