

Genomic Databases and Genomic Privacy: Can we have the best of both worlds?

Widespread scientific curiosity and the possibility of medical benefits make it nearly impossible to justify keeping genomic data under lock and key. At the same time, there has been increased concern about the private information leaked by de-identified genomic databases. What is the right middle ground?

In the first half of this talk, I will present some ongoing work investigating privacy risks in a publicly available genomic database, and possible approaches to help mitigate these concerns. The second half of the talk will focus on how ideas from Bayesian statistics and the statistical disclosure control literature might help turn these mitigation techniques into quantitative guarantees.