Review: Understanding Database Reconstruction Attacks on Public Data

I applaud you all on the writing. Very clear and easy to follow!

In the abstract you argue that SDL methods in general cannot prevent reconstruction attacks, but in the body of the paper you focus on just one SDL method, cell suppression. I think you should be more specific in the abstract about cell suppression because I think that there are SDL methods that might provide more protection. For example, I think that noise addition itself is considered an SDL method.

In the abstract and at the end of Section 1, you say that you are going to discuss the implication for the decennial census. Maybe that was taken out? I don’t actually see that discussion in the later sections.

My only real critique is that you don’t make a connection between the hypothetical example in Section 2 and real-life data. You’ve demonstrated that the data in Table 1 can be reconstructed, but how similar is Table 1 to something that has actually been released by a statistical agency in the past? Even if you don’t try to reconstruct such a real-life example, I think it would good to include one so that you can make the argument that such data has been released publicly in the past.

I noticed a few typos in my reading. I’ll put my paper copy with those noted under your door Simson.