

COMP8400 – Tutorial 4 questions
Privacy-preserving data mining (and text mining)

Q1: In the introduction of their paper on privacy-preserving data mining, the authors write about *sensitive raw data* (and provide several examples), as well as *sensitive knowledge*. Give three examples of sensitive knowledge.

Q2: Describe one example application each, where data would be *horizontally* or *vertically* distributed.

Q3: What are the two metrics used in the paper to quantify the degradation in database performance when data is being modified?

Q4: What are the three main types of privacy-preserving data mining techniques?

Q5: What are the main evaluation criteria for assessing privacy-preserving data mining algorithms and tools?

Q6: What is the main conclusion of this paper?

Questions / issues you like to have discussed in this tutorial: