

Exercise D — Tool triage

(Leveraging) AI for Research Purposes · Class 3 · 20 October 2026

For each task: **what you ask the tool to do** (compress, extract, expand, formalise, search) → **tool family** → **one concrete tool** → **the main risk** → **the level of the data**, and **whether your tool is allowed there**.

Task	Use	Tool family	Concrete tool	Main risk	Data level · allowed?
D1					
D2					
D3					
D4					
D5					
D6					

The six tasks

D1. You are starting a new paper on a question nobody in your group has worked on. You want to know what has been published in the last five years. *Data: public.*

D2. You are doing a meta-analysis and need sample sizes, effect sizes and confidence intervals from 40 published PDFs. *Data: public.*

D3. You have 200 excerpts from interviews you conducted for your own study, with consent for research use. You want a first pass of codes before coding by hand. *Data: confidential; secret if the interviews were part of a clinical study.*

D4. Your survey platform exports 3,000 rows of pseudonymised answers, with inconsistent date formats, free-text in numeric fields and three different spellings of every canton. You need a clean dataset and you are not good at R. *Data: confidential.*

D5. You have 30 hours of recorded interviews with nurses about their working conditions, for a study on staff shortages. You need transcripts. *Data: confidential.*

D6. Your manuscript is ready to submit and has 74 references, some added in a hurry. You want to be sure every one exists and is cited correctly. *Data: the references are public; the manuscript is internal.*