Team Metadata. Writeup of Day 1.

Goal: Identify Case vs Control samples.

Map run data (runs provided for hackathon) to ontology terms and intersect this data set with metaSRA data which has normalized metadata terms.

Initially only 9K of the hackathon runs overlapped with metaSRA data. These have been exported for summary statistics.

Still under investigation is why the remaining 14K samples didn’t map to metaSRA.

We have produced a fake data set of run numbers mapped to “terms” and placed in “studies”. This data set is being used to create the query script to enable searching of “case” vs “control” for the 9K runs in both hackathon set and found in metaSRA.