

Instructions for using judge and federal case data

The “master” data for federal judges is a CSV-formatted file, with a header row naming the fields. The first value in each row is the ***judge id***. The file can be found at:
<https://s3.us-east-2.amazonaws.com/lexisai-hackathon/entity-masters/federal-judges.csv>

The cases you’ll be using are in the following directory:
<https://s3.us-east-2.amazonaws.com/lexisai-hackathon/harvard-case-law/fcass/casemets/>

The file names of the cases contain the ***case id***, separated by the string “_redacted_CASEMETS”. For example, case id **32044046873683_0164** can be found in the following file:
32044046873683_redacted_CASEMETS_0164.xml

To map cases to judges, use the following file:
<https://s3.us-east-2.amazonaws.com/lexisai-hackathon/harvard-case-law/fcass/caseld-judgeld.tsv>

The file contains tab separated case id and judge id, e.g.
32044046873683_0164 154

This means judge 154 was identified for case 32044046873683_0164.