本次数据收录了目前标注完毕的串案判决书信息，所有判决书均由三个人进行了标注，且每起串案（在多份判决书中出现相同的原告与涉案专利）至少包含两份判决书。共计64起串案，200份判决书，用于明细比对。

由于本次数据规模较小，暂不划分训练集与测试集，由于每起串案至少有两份判决书，可以采取以其中之一与其他所有相关案件进行比对的方式。

数据格式为json格式。

data ——串案1

——caid：该串案的唯一编号

——docs：包含的判决书

——判决书1

——id：该判决书的唯一编号

——description：判决书描述，目前暂用编号代替

——content：判决书内容，以短句为划分

——answer：标注结果

——标注1

——index：标注的句子位置，对应content中的短句下标，位置从1开始

——original\_tag：原始的标注结果

——用户1的标注:

——username：用户名

——value：标注结果，格式为“一级标签/二级标签/...”，可能有多个

——用户2的标注：

——用户3的标注：

——tag:按照标签整理后的标注结果（不考虑标签的层级关系）

——标签1：

——value：该标签名

——weight：标注该标签的人数

——标签2：

——标签3：

......

——标注2

——标注3

——标注4

......

——判决书2

——判决书3

......

——串案2

——串案3

......