中国人民法院网爬虫代码说明：

analysis：处理并解析下载的json数据，提取所需数据，再以约定好的json格式输出。

analysis\_bankruptcy：抓取破产文书的分析器，增加咯了pdf下载链接url的数据提取。

generator：生成含json数据的目标页url

generatorByDay：每日更新生成器，每次运行时判断pickle时间与当前时间是否一致，一致则使用上次的索引，不一致则重置索引。生成url后解析json数据，判断日期是否于当日一致，一致则输出url，不一致则不输出。（每日生成器代码是蔡丹红改写的，具体逻辑可以问问她）

generator\_bankruptcy：单独生成破产文书目标url，代码只做了相应url构造部分的修改。

generator\_bankruptcy\_by\_day：每日生成更新的破产文书新增条目url。