

## 一、 人名提取设计方案：

模拟接受语音输入并模拟语音转文字

输入： 包含人名的文字

算法流程：

利用预定义模版匹配人名；

if 匹配失败：

    利用stanford corenlp提取人名

利用python工具包将人名转化为拼音；

基于拼音，与预定义人名进行模糊匹配，包括前后鼻音/ln等常见发音错误；

返回基于拼音的匹配结果。

if 匹配成功：

    返回人名

## 二、 请假理由提取方案：

需要的字段： 时间、审批人、理由、请假类型

例如： 下周一到周五请个假，让xxx帮我批一下，我要出去玩

算法流程：

分割每一句话

对于每一句话：

    如果包含审批人或是请假类型：

        去掉这句话

    continue

如果该句子包含请假：

    去除该句包含的“请假”动词词组

    如果有其他动词：

        添加到理由中

剩下所有句子都添加到理由中