

法律大数据与人工智能应用

马志柔 亢良伊 刘杰 叶丹

联系方式 (13488689166, ljie@otcaix.iscas.ac.cn)

研究自然语言处理、机器学习和知识图谱技术，高效的从法律文书中提取法律术语特征、构建法律知识图谱，探索法律检索、文件审阅、案件预测、咨询服务等模型，提供智能化的法律服务。

关键技术

1: 基于机器学习的实体识别技术

技术：将文本信息的识别与抽取转换为序列标注和分类问题，采用CNN + Bi-LSTM + CRF相结合的方法训练语言模型，结合概率模型和深度学习两者的优点，既考虑了词与词之间的语义关系，又考虑了词在整句话中的重要性。

效果：该网络结构实现了端到端模型，可以更大范围应用到不同领域的序列标注任务。

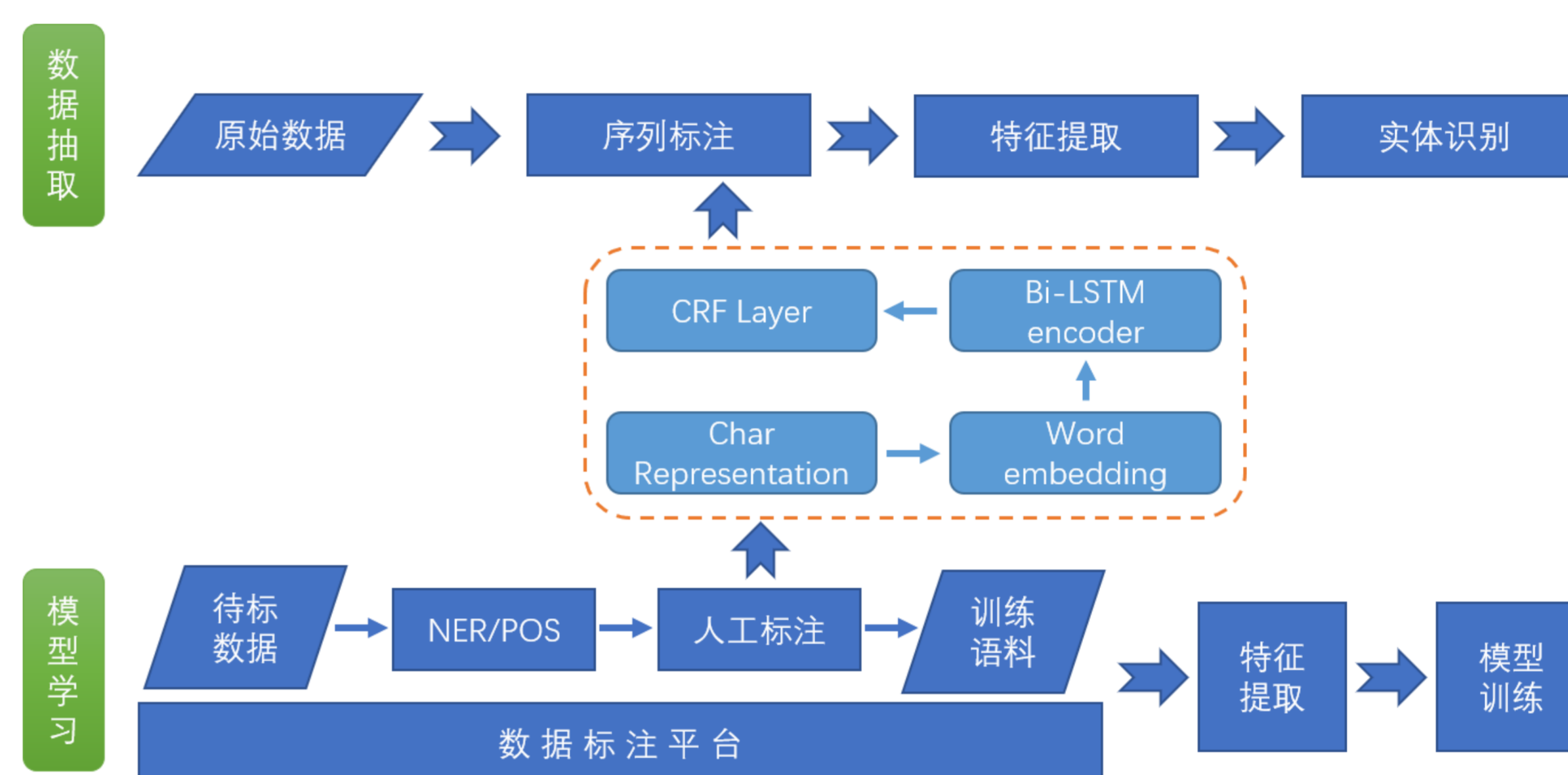


图1 实体识别技术架构

2: 基于知识图谱的智能问答技术

技术：针对法律领域的特殊性，通过层次化标签和推理规则集来刻画法律的知识类型和表示模型，以一种松散耦合的关系来构建法律领域知识图谱；利用CNN/RNN + Attention方法实现对短文本问题的端到端的意图分析理解。

效果：该问答模型既考虑问题的理解分析，又考虑法规的推理过程，实现了情景化问答。

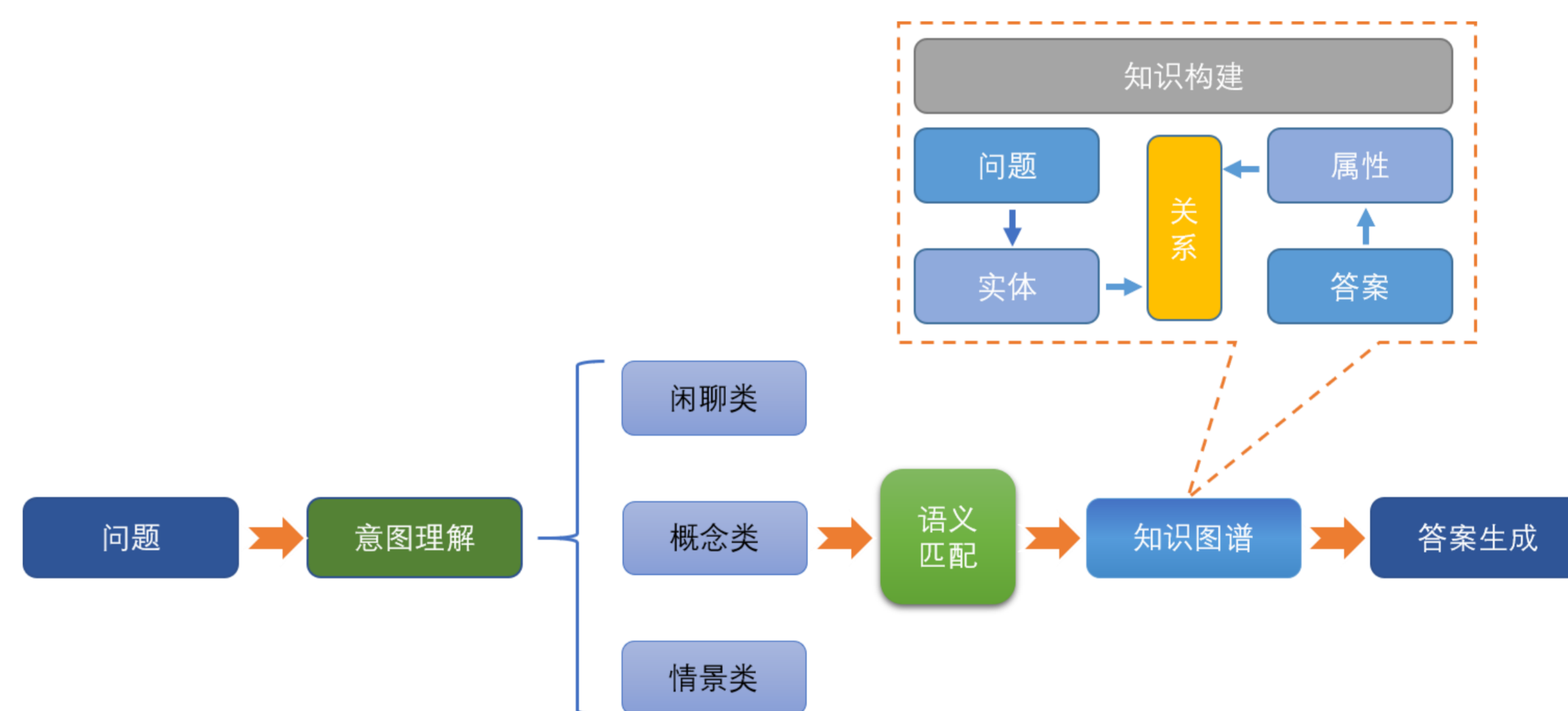


图2 智能问答技术架构

应用案例

与无讼网络科技有限公司联合成立“法律人工智能实验室”开展两项应用

智能预判系统

针对交通事故类案件，通过标准判别、证据提取、规则推理，形成一个赔偿标准的“智能法官”，为保险人员提供决策服务。



图3 平安保险车险人伤理赔“智能法官”

智能问答系统

针对劳动领域类问题，通过问答理解、语义匹配、答案生成，形成一个法律解答的“律师朋友”，为法务人员提供咨询服务。



图4 无讼法小淘智能解答“律师朋友”