



DeepLaw

专属深度法律顾问

问题背景



法律现状



案件 多

2015年全国法院新收案件近 1800万 件，同比上升 22.81%

法制网 2016.03.18



法律 细

我国现行法律法规中，共有 29 类大大小小共计 2569 部

中国企业集成网数据



咨询 难

我国执业律师人数已超过 29.7万 人，律师事务所达到 2.4万 多家

新华网 2016.03.30

法律咨询需求 与日剧增，却不知 找谁帮忙、从何问起，而且 费用高、耗时多

应用场景

应用场景

DeepLaw

专注于 **法律** 垂直领域的人工智能聊天机器人



随便唠唠

一般聊天机器人所具备的
日常交流能力
可调侃、可逗趣、可玩耍



专业问答

基于法律法规和法律术语的
专业知识问答
可学习、可解惑、可答疑

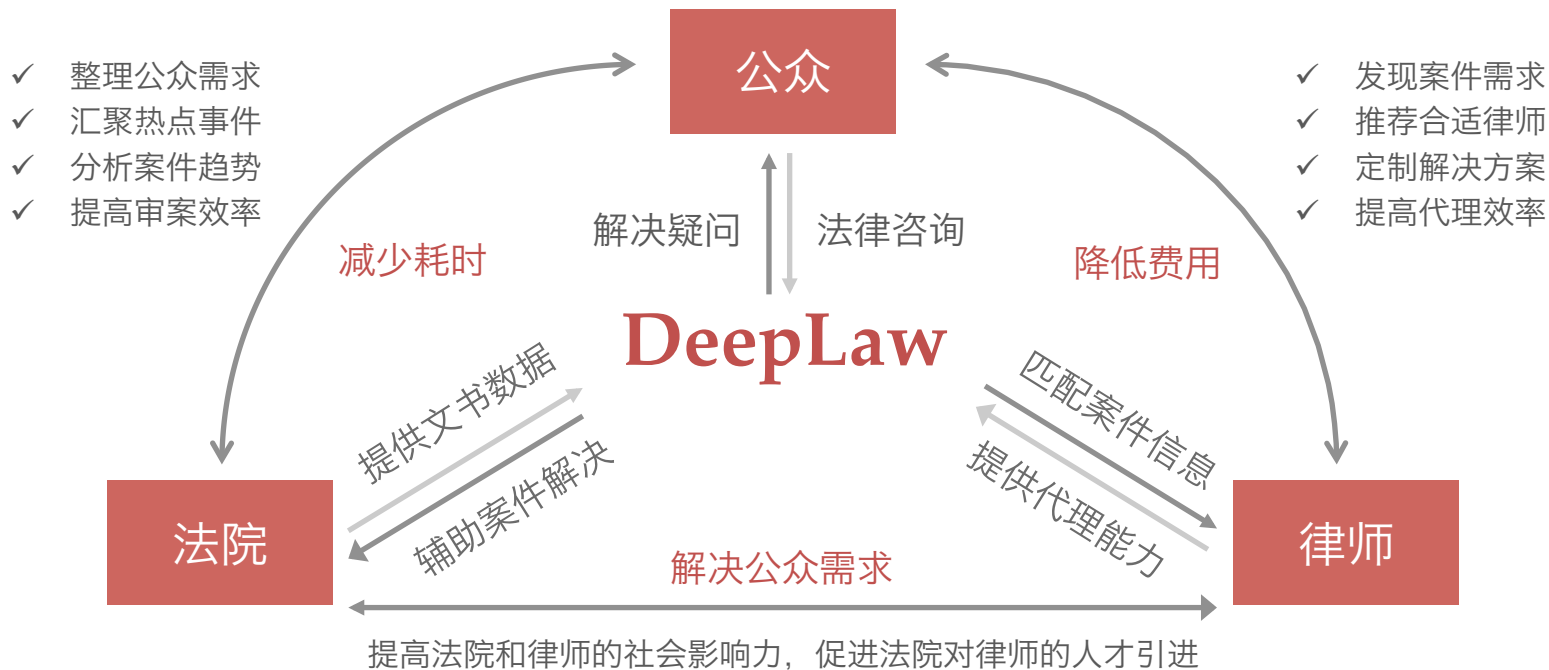


法律咨询

类自然语言和自由定制的
法律信息咨询
可谈心、可讨论、可交流

商业模式与价值

商业模式与价值



理论框架



理论框架

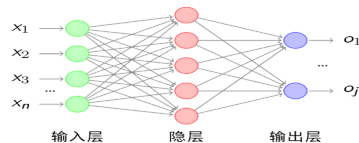
1 数据层

聊天语料 问答数据 领域本体 知识图谱



2 模型层

关键词搜索 模板匹配 机器学习 神经网络



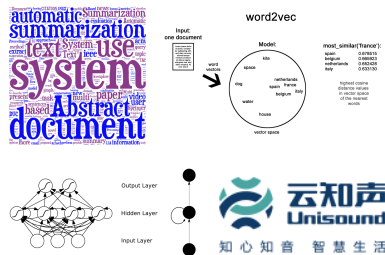
用户输入

对话输出

3 理解层

句法分析 语义理解

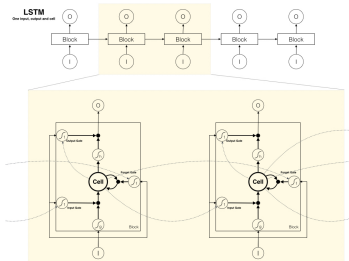
语境感知 语用计算



4 语言层

文本搜索 知识匹配

抽取式 生成式



数据层

语料和数据主要来自于 互联网抓取

聊天语料

基于 字幕库 提供的电影和电视剧 对话字幕数据 获取聊天语料



问答数据

基于 法律咨询、华律网、法律快车 等网站收集已解决的咨询问题 和相应的 答案



领域本体

基于 中国法院网 采集专业 法律术语 和各类案由所涉及的 本体



知识图谱

基于 搜狗百科、百度百科、互动百科 法律类别下的词条构建法律垂直领域下的 知识图谱



模型层

关键词搜索

提取输入中的关键词
并在语料库中搜索相
关的对话



模板匹配

基于预先定义好的文
本模版在语料库中匹
配相关的对话



机器学习

基于机器学习模型对
聊天数据进行训练

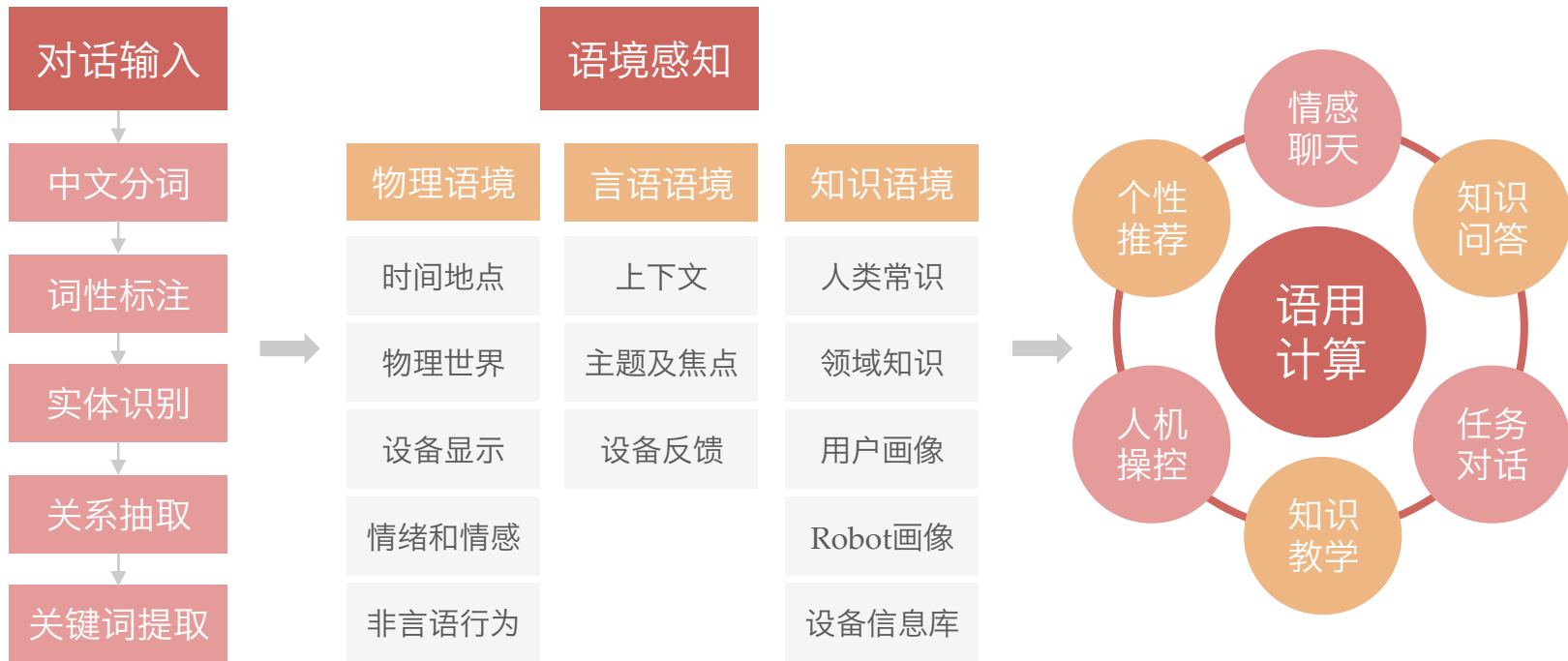


神经网络

基于LSTM等递归神
经网络模型对聊天数
据进行训练



理解层



语言层



技术实现



技术实现

语言层

DeepLaw



基于 Flask 实现产品交互界面

理解层

基于 Keras 实现深度学习



Keras

使用 scikit-learn 进行机器学习



模型层



jieba
word2vec

基于 Jieba分词 进行初步文本处理

使用 Word2Vec 训练语言模型

数据层



基于 Scrapy 和 Redis 获取海量
多源异构数据并存储至 MySQL

对话

应用

模型

信息

语料

数据流

使用流程



产品界面—开发中



- ✓ 用户 **手动指定** 模式，随便唠唠、专业问答或法律咨询，DeepLaw 以相应模式进行聊天
- ✓ 用户 **直接输入** 聊天内容，DeepLaw 根据聊天语境 **自动判断** 最优模式，并进行聊天



DeepLaw

下一代人工智能法律顾问