

基于星火大模型 API 的法律诉状书生成系统的技术文档

版本： 0.0.1

成员： 曾嘉怡（队长），范珑骁，刘文涛，王仁淞，张逸飞，
王凯文，李兵磊（排序按照成员进群先后顺序）

队伍： 张三李四队

日期： 2023 年 8 月 24 日星期四

目录

1. 引言	4
1.1.目的.....	4
1.2.范围.....	4
1.3.定义、缩略语和缩写	5
1.4.参考文档	5
2. 系统概述	6
2.1.总体描述	6
2.2.系统架构	7
3. 详细设计	7
3.1.模块描述	7
3.2.数据结构设计	8
3.3.错误输入处理	8
3.4.流程设计的约束.....	9
3.5.技术细节	9
3.5.1.整体框架技术	9
3.5.2.模糊匹配技术判断用户输入.....	10
4. Prompt 设计	10
4.1.设计 Prompt 充分调用大模型的原子能力	10
4.2.设计 Prompt 严格规范大模型输出格式.....	11
4.2.1.限制大模型以标准数据结构输出（json 格式）	11
4.2.2.限制大模型按照标准模板形式输出	11
4.2.3.限制大模型语言规范.....	11
4.2.4.让大模型在原先已有内容的基础上追加新内容。	12
5.安装和配置指南	12
6.使用指南	12
6.1.用户界面与功能概述	12
6.2.常见操作	14
7.测试	15
7.1.测试策略	15
7.2.测试用例	15
7.2.1.测试用例 1	15
7.2.2.测试用例 2	16
7.3.测试结果	17
7.4.测试结果分析	18

8.思考与反思	19
8.1.用户交互体验设计	19
8.2.大模型在特定任务下的灵活性和可控性	19
8.3.Langchain 技术的应用	19
9.附录	20

1. 引言

1.1. 目的

本技术文档旨在详细介绍基于星火大模型 API 的法律诉状书生成系统。该系统利用先进的机器学习技术，通过自动化分析和理解，为法律专业人士和相关人员生成精确、高质量的法律诉状书。

主要受众包括：

- 开发人员：本文档将辅助开发团队理解如何整合星火大模型 API，实现系统的架构和功能。
- 法律专业人士：系统的最终用户，通过本文档的使用指南部分，将了解如何操作系统以生成所需的法律文书。
- 项目管理和利益相关方：通过全面了解系统设计和功能，确保项目符合期望的目标和质量标准。
- 测试和维护团队：本文档将提供足够的技术细节，支持系统的测试、维护和可能的未来升级。

本项目结合了法律领域的专业知识和先进的人工智能技术，旨在提高诉状书编写的效率和准确性，减轻法律工作者的工作负担，同时确保文档的专业性和合规性。

文档涵盖了从设计理念到实际操作的全方位内容，为项目的各个阶段和各个层面的参与者提供支持和指导。

1.2. 范围

本技术文档的适用范围涵盖了基于星火大模型 API 的法律诉状书生成系统的全方位内容，具体包括：

- 系统架构：描述系统的整体设计和组件结构。
- 功能说明：详细解释系统的各项功能，包括诉状书生成、模板管理、用户界面操作等。
- 接口描述：包括与星火大模型 API 的交互以及其他可能的系统集成。
- 安装和配置：提供系统部署在不同环境中的指导方针。

5. 使用指南：针对不同用户角色，如法律专业人士、系统管理员等，提供操作教程。
6. 测试策略：包括测试计划、测试用例和预期结果等。
7. 维护和支持：描述系统的日常维护、故障排除和升级路径。

请注意，本文档主要集中在技术方面，涉及系统的设计、开发、部署和维护等。一些与项目管理、成本评估、市场分析等商业方面的细节可能不在本文档的讨论范围内。此外，虽然文档将详述如何使用星火大模型 API，但不包括该 API 自身的详细技术文档或底层机器学习算法的深入解释。

1.3.定义、缩略语和缩写

以下是基于星火大模型 API 的法律诉状书生成系统技术文档中可能使用的一些特殊术语、缩略语和缩写的列表：

1. API（应用程序编程接口）：一组定义如何通过应用程序访问某些功能的规程和协议。
2. 星火大模型：本文档中指的特定机器学习模型，用于支持诉状书的自动生成。
3. 诉状书：法律文件，用于提起法律诉讼。
4. UI（用户界面）：系统的图形界面，通过它用户与系统进行交互。
5. DB（数据库）：用于存储系统中的数据和模板等信息的存储库。
6. ML（机器学习）：一种人工智能技术，用于训练模型从数据中学习。
7. NLP（自然语言处理）：与机器学习有关的领域，专注于让计算机理解和解释人类语言。
8. HTTPS（超文本传输安全协议）：用于安全通信的协议。
9. OS（操作系统）：支持系统运行的底层软件环境。

1.4.参考文档

Streamlit: [Streamlit documentation](#)

星火 API: [星火认知大模型 Web 文档 | 讯飞开放平台文档中心 \(xifyun.cn\)](#)

2. 系统概述

2.1. 总体描述

基于星火大模型 API 的法律诉状书生成系统是一项创新的技术解决方案，结合了先进的人工智能技术和法律专业知识，为法律专业人士和相关团体提供自动化的法律诉状书生成服务。

1. 原告被告信息提取	2. 案由类别确认	3. 诉讼请求生成	4. 事实与理由生成	5. 起诉状生成
<ul style="list-style-type: none">允许输入多个原告与被告的详细信息支持自然人和非自然人信息的输入	<ul style="list-style-type: none">允许用户自行选择支持案由智能识别	<ul style="list-style-type: none">根据案由智能生成诉讼请求根据用户输入自动完善诉讼请求	<ul style="list-style-type: none">根据案由智能生成事实与理由提示根据用户输入自动完善事实与理由	<ul style="list-style-type: none">预览修改保持起诉状

- 原告被告信息提取模块：该模块允许用户输入多个原告和被告的详细信息，包括自然人和非自然人的信息。通过用户提供的输入，系统将自动提取和记录相关信息，以备后续使用。
- 案由类别确认模块：用户可以在此模块中选择案由类别，也可以选择让系统自动智能识别案由。系统支持多种案由类别，并根据用户选择或智能识别的结果，为后续步骤提供指导。
- 诉讼请求生成模块：在该模块中，系统根据选定的案由智能生成初步的诉讼请求，包括一般性要求等。用户还可以根据需要自行输入补充信息，以使诉讼请求更加准确和完整。
- 事实与理由生成模块：根据已确认的案由，系统将智能生成相关的事实与理由提示。这些提示有助于用户更好地组织和阐述案件的事实和法律理由。用户可以基于系统生成的提示进行进一步补充和修改。
- 起诉状生成模块：在该模块中，用户可以预览并修改由系统生成的起诉状。系统将根据之前输入和生成的信息，自动生成一个初步的起诉状，用户可以对其进行修改和调整，直到满意为止。

通过以上模块的协同工作，该系统能够帮助用户高效地提取原告被告信息，确认案由类别，生成诉讼请求和事实与理由，并最终生成可供预览和修改的起诉状。这些功能的结合为用户提供了一个完整且方便的法律辅助工具，以提高法律文书的准确性和效率。

2.2.系统架构

基于星火大模型 API 的法律诉状书生成系统是一个结构复杂的应用，包括了前端、后端和语音合成等组件。以下是系统的架构描述：整个架构确保了系统的灵活性和可扩展性，同时保持了高质量的输出和良好的用户体验。

1. 前端界面（Streamlit）： 提供用户交互的界面，用户可以通过这里选择模板、输入数据并查看生成的诉状书。
2. 后端服务（星火大模型 2.0 API）： 负责处理用户的输入，调用星火大模型进行文本生成，确保输出的诉状书准确、合规。
3. 语音合成服务（星火大模型语音合成（流式版）API）： 将生成的诉状书文本转化为语音，支持多种场景和需求，例如盲人用户的访问或语音助理的集成等。

3. 详细设计

3.1.模块描述

按设计 UI 分类的话分为 4 大块：用户聊天界面、预览诉讼文书、法律知识汇总、查看历史输入。

用户聊天界面：提供一个交互式的聊天环境，让用户可以直接通过自然语言向系统提出请求和问题。文本/语音输入，用户可以通过键入文本或者语音与系统进行交流。智能响应，基于星火大模型 API 生成的回复，实时与用户对话。上下文保留，对话历史记录保持，确保连贯的交流体验。

预览诉讼文书：让用户可以预览和审查生成的诉讼文书，确保准确无误。实时预览，文书生成后，用户可以立即查看完整内容。编辑和修改，提供编辑工具，用户可以对生成的文书进行个性化修改。导出选项，支持不同格式的导出，如 PDF、Word 等。

法律知识汇总：集成和展示与诉讼文书相关的法律知识和参考资料。法律条款查询，提供搜索和浏览功能，让用户可以查找特定法律条款和规定。相关案例，展示与当前诉讼主题相关的法律案例和判决。帮助和指导，提供有关如何使用系统和生成文书的帮助和指导材料。

查看历史输入：让用户可以访问和管理先前使用系统时的历史记录和输入。历史记录浏览，列出用户过去的操作和输入记录，按日期和主题排序。重用和编辑，允许用户重新使用过去的输入，作为新诉讼文书的基础。数据保密，确保历史记录的安全存储和访问控

制。

3.2.数据结构设计

为了更好地存储和管理生成法律诉讼文书所需的信息，我们对数据结构进行设计。根据法律文书的具体内容，json 文件中主要存储如下信息：

1. 原告和被告信息：将原告和被告的个人信息以更详细的方式进行存储，为了扩展性考虑，支持多个原告和被告的情况。
2. 法定代理人和委托诉讼代理人：为代理人信息添加律师事务所名称、联系方式等。
3. 证据信息：存储有关证据的更详细信息。包括证据名称、证据来源、证据描述等。
4. 日期格式：使用系统日期以便更方便地进行日期的比较和计算。
5. 扩展性考虑：将"诉讼请求"和"事实理由"字段设计为多行文本，可以容纳更长的内容。

```
{
  "原告": [{person_info01},{person_info02}],
  "被告": [{
    "姓名": "张**",
    "性别": "男",
    "出生日期": "1980 年 3 月 6 日",
    "民族": "汉族",
    "住址": "甘肃省平凉市静宁县",
    "联系方式": "135791****",
    "身份证号": "320431*****",
    "法定代理人": "王**",
    "委托诉讼代理人": {"姓名": "***", "事务所": "***事务所"}
  }],
  "案由": "房屋买卖合同纠纷",
  "诉讼请求": "诉讼请求相关内容",
  "事实理由": "事实理由相关内容",
  "证据": "证据相关内容",
  "法院": "对应法院",
  "日期": "2023 年 08 月 23 日"
}
```

3.3.错误输入处理

用户输入针对 prompt 的回答可以从完整程度分为回答完整和回答不完整，从相关程度可以分为内容相关和内容不相关。以下是四种情况的分别举例：

例如针对给出的引导提问：谢谢您提供的信息！请告诉我 自然人（个人） 的如下信息、姓名、身份证号、性别、出生日期、民族、住址、联系方式、委托代理人（如有）、法定代理人（如有）。

完整相关回答：我叫王大锤，男性，身份证号码是 410224199908025735，出生日期是 1999 年 8 月 2 日，民族汉，住在上海市杨浦区，手机号码是 13586932058，我没有委

托代理人和法定代理人。

不完整相关回答：我叫王大锤，男性，住在上海市杨浦区。

不相关回答：我很喜欢唱歌，我今晚想在人民广场路演唱歌。

针对这三个不同的回答，我们采取不同的策略，针对 **1.完整相关回答**，我们提取所需要的相关信息，继续下一步的引导。针对 **2.不完整回答**需要给出用户缺失的信息，继续引导用户回答。针对 **3.不相关回答**，提供委婉提示，引导用户回归正确信息输入方向。

3.4.流程设计的约束

我们本着以人为本，用户至上的设计理念。为用户提供尽可能清晰，友好，温馨的对话界面，为了方便用户输入，我们设计语音输入为主，文本输入和纠正为辅的输入模式。同时严格保护用户的隐私和信息。总结起来分为以下四个方面：

1. 数据保护和隐私：我们承诺严格遵守相关的数据保护法律法规，确保用户输入的信息得到适当的保护和处理。所有数据将仅用于提供服务，并在不再需要时进行安全删除。
2. 引导用户输入：为了确保用户能够提供准确和详尽的信息，我们将根据不同的对话情境，逐步引导用户提供必要的信息。通过明确的问题或提示，帮助用户理解所需信息内容。
3. 错误处理机制：考虑到用户可能会输入错误或不确定的信息，我们将设置错误处理机制。通过提示用户重新输入或提供纠正选项，确保用户能够及时修正错误并继续进行对话。
4. 文书预览修改：为了保证生成的文书准确无误，我们将提供文书预览修改功能。用户可以在预览界面中查看生成的文书内容，并对其进行修改和调整，以满足自身需求。

3.5.技术细节

3.5.1.整体框架技术

在这个项目中，我们使用了 Streamlit 框架来构建用户界面。Streamlit 是一个用于创建数据应用程序的 Python 库，它提供了一种简单易用的方法，能够快速地将数据可视化和机器学习模型部署到网页上。

在用户界面中，我们采用了嵌入的 JS 代码来实现语音输入功能。通过调用浏览器的 Web Speech API，我们可以使用户能够通过麦克风输入语音，并将其转换为文本。这样，用户就可以通过语音来与应用程序进行交互，提供更加便捷和自然的输入方式。

另外，我们使用了星火语音合成 API 来实现语音合成功能。星火语音合成 API 是一个

强大的语音合成服务，它能够将文字转换为自然流畅的语音输出。通过将应用程序生成的文本信息传递给星火语音合成 API，我们可以获得符合语音合成需求的音频文件，并将其播放给用户。

3.5.2.模糊匹配技术判断用户输入

使用模糊匹配技术判断用户回答结果与预期结果的一致性。

获取用户输入的回答结果和预期结果，并进行预处理。预处理可以包括去除标点符号、转换为小写字母等操作，以便更好地进行比较。使用字符串比较算法（如 Levenshtein 距离）计算用户回答和预期结果之间的相似度或距离。根据设定的阈值，判断用户回答是否与预期结果相似。根据判断结果做出相应的处理。如果匹配成功，则可以给予用户正确的反馈；如果匹配失败，则可以提示用户重新回答或提供更准确的答案。

4. Prompt 设计

4.1.设计 Prompt 充分调用大模型的原子能力

1、依靠大模型对国家各个省市区的了解，让其自动补全地址信息

Prompt:（用户输入地址）请帮我补全以上住址使其包含省市区信息，并以如下 json 格式输出:{"住址":""}

2、依靠大模型对各地法院的了解，让其根据被告人住址自动生成可以上诉的法院名称

Prompt: 以下是被告人的住址，请列举他们住所地的人民法院，并只输出法院名称。
住址是（被告人住址）

3、依靠大模型对案由类型的理解，让其自动判断用户描述案件的案由类别

Prompt: 请根据用户输入判断案件类型，案件类型一般有机动车交通事故责任纠纷、民间借贷纠纷、离婚纠纷、合同纠纷、买卖合同纠纷、金融借款合同纠纷、借款合同纠纷、劳动争议、房屋买卖合同纠纷、建设工程施工合同纠纷、土地经营权纠纷等，如果不是以上案件类型，请自行判断，请注意一定要以以下 json 格式输出案件类别{"案由":""}，用户的输入是：（案情描述）

4、依靠大模型对各个案由的理解，让其自动生成该类型案由可能的诉讼请求和事实和理由

Prompt: 案件类型是（案件类型），请给出在起诉书中典型的该类别案件的诉讼请求，并分点陈述

Prompt: 案件类型是（案件类型），请给出在起诉书中典型的该类别案件的事实与理

由，并分点陈述

4.2.设计 Prompt 严格规范大模型输出格式

4.2.1.限制大模型以标准数据结构输出（json 格式）

为使得大模型输出的内容能够稳定地被程序解析，我们设计 prompt 让大模型提取个人、公司信息后将数据转化为 json 格式输出，prompt 设计样例如下：

1) 判断是否有委托代理人，若有则填入个人信息的 json 文件中

Prompt: 客户将告诉你是否需要委托代理人。若用户没有或者不需要委托代理人，则不变；如有请填写下面的 json 模板中，{'姓名': null, '性别': null, '出生日期': null, '民族': null, '住址': null, '联系方式': null, '身份证号': null, "法定代理人": null, '委托诉讼代理人': {'姓名': null, '事务所': null}}

2) 提取客户输入的个人信息并填入 json 文件中

Prompt: 客户将给你一段文字，需要你提取全部关键信息，不丢失信息，填充到下面的 json 文件中，并符合 json 语法，记住只需把这个 json 文件输出给我。{'姓名': null, '性别': null, '出生日期': null, '民族': null, '住址': null, '联系方式': null, '身份证号': null, "法定代理人": null, '委托诉讼代理人': {'姓名': null, '事务所': null}}。提问和回答如下：（用户输入的个人信息）

3) 将补全后的住址填入 json 文件中

Prompt: （用户输入地址）请帮我补全以上住址使其包含省市区信息，并以如下 json 格式输出：{'住址': ""}

以上输出数据格式规范使得程序在处理大模型输出时十分稳定，几乎不会出错。

4.2.2.限制大模型按照标准模板形式输出

诉讼书里的格式非常规范，为了让大模型生成更规范的内容而不被用户输入的内容带偏，将之前生成或者提前预备的模板加在 prompt 前面作为参考。

Prompt: （模板）以上是可以参考的诉讼请求模板，请根据以下原告描述内容生成诉讼请求部分，注意严格按照原告意思，不要随意篡改，不要出现第一第二第三人称，请注意分点阐述。原告：（诉讼请求）

4.2.3.限制大模型语言规范

大模型在生成答案时不稳定，如果不在 prompt 中加以限制，有可能会生成随机的或

者是编造的内容，这与我们严谨的流程不符，所以我们对大模型生成的文本加以各种类型的限制，来尽可能让其以我们想要的格式生成合规的、客观的内容。

1) 要求大模型生成诉讼请求提示，但是限制案件类别，并要求分点阐述。

Prompt: 案件类型是“房屋买卖合同纠纷”，请给出在起诉书中典型的该类别案件的诉讼请求，并分点陈述。

2) 按模板生成诉讼请求/事实理由，但是要求大模型不得篡改原意，不能出现人称代词，分点阐述等限制。

Prompt: (模板)。以上是可以参考的诉讼请求模板，请根据以下原告描述内容生成诉讼请求部分，注意严格按照原告意思，不要随意篡改，不要出现第一第二第三人称。请注意分点阐述。(原告描述内容)

3) 限制输出的内容本身，要求只需要输出名称。

Prompt: 以下是被告人的住址，请列举他们住所地的人民法院，并只输出法院名称。住址是(住址)。

4.2.4.让大模型在原先已有内容的基础上追加新内容。

这类 Prompt 的产生主要是建立在用户补充信息这一环节的基础上，原信息格式已经固定，我们只需要大模型在原文中追加新内容，所以需要限制模型的输出格式，防止大模型追加格式与原本格式不符。

补充诉讼请求/事实与理由时，限制模型将新信息整理到原文中，并间接限定了输出格式，并限制了引导词的使用。

Prompt: 给定原先的“XXXX”部分(原 XXXX 部分内容)现有新的未确定类型的信息需要添加到原先内容中。原先的内容如上所述，新的信息如下：**(新信息)**。请根据新的信息的性质和内容，整理到原先的“XXXX”中，然后输出“XXXX”，不需要有引导词

5.安装和配置指南

我们提供了运行代码所需要的环境配置参数，可以很简单的在本地或云端运行起来。详情见配置文档！

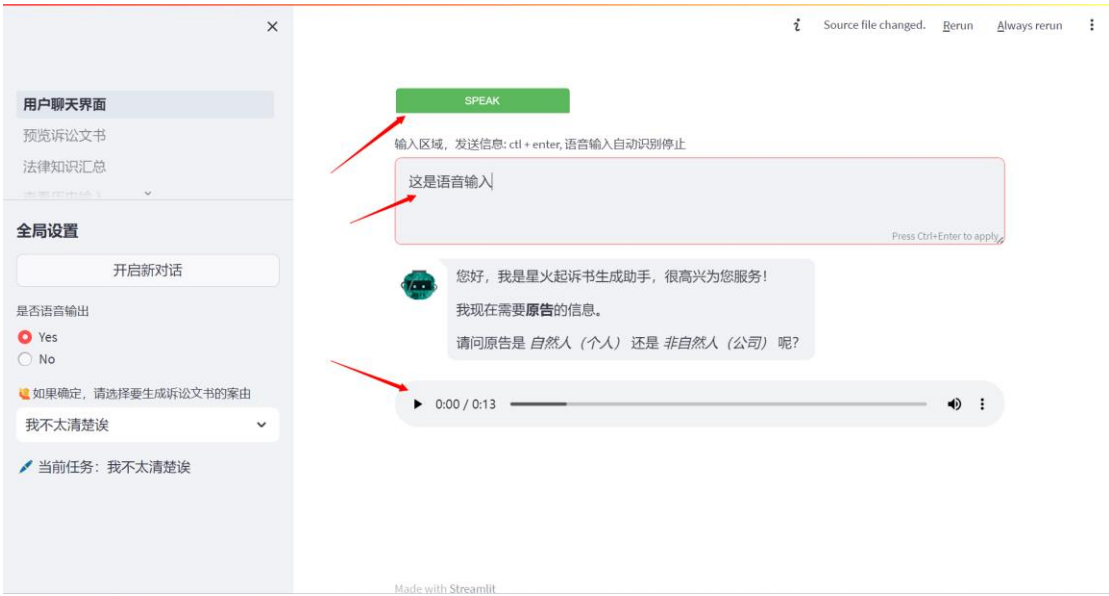
6.使用指南

6.1.用户界面与功能概述

我们的模型基于星火大模型 API，集成了先进的人工智能技术与专业的法律知识，完成一套用户交互友好的法律诉状书生成系统。该系统基于支持语音输入的交互式聊天界

面，集诉讼文书生成、文书预览和编辑，知识库访问，历史记录管理等功能为一体。其主要的用户界面和界面功能如下：

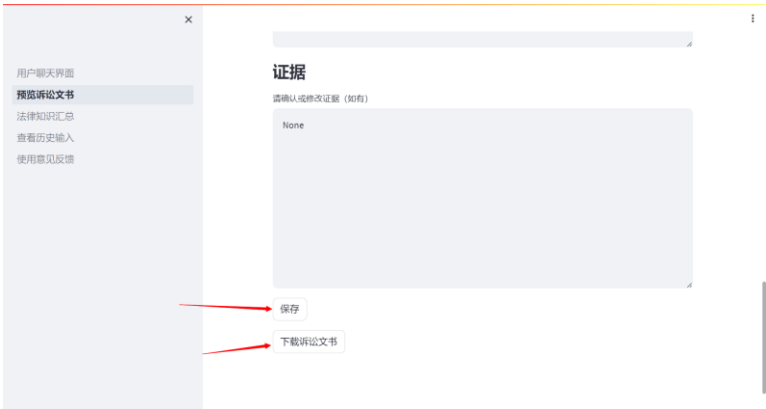
下图是用户聊天界面，如图所示，系统可以接受语音的输入输出，并且会引导用户输入信息。



下图是用户输入完成之后，预览诉讼文书界面，在此界面用户可以检查之前的输入是否有错误和遗漏，用户可以在此界面进行修改和补充



下图是用户确认无误之后，用户可以选择保存并下载 docx 文件。



下图是法律汇总，系统可以根据用户的输入，进行相关科普和推荐案例。



下图展示了系统的历史聊天数据查询，可供用户查询和回看



6.2. 常见操作

以下是用户可能需要执行的一些常见操作，以及相关的教程和提示：

1. 生成诉讼文书
 - a) 登录系统并选择“用户聊天界面”选项。
 - b) 通过聊天界面输入相关案件和诉讼信息。
 - c) 选择适当的法律条款和模板。
 - d) 预览生成的文书。
 - e) 如有需要，编辑文书并保存或导出。
2. 查阅法律知识
 - a) 在“法律知识汇总”模块中查看相关的关键词或条款。
3. 管理历史记录

a) 选择“查看历史输入”模块。

b) 浏览先前的操作和输入记录。

4. 使用语音合成

a) 在预览诉讼文书界面，选择“语音播放”选项。

b) 点击播放，通过音响设备收听生成的语音。

这些操作指导旨在帮助用户轻松、高效地利用系统的功能，实现其法律诉状书生成的需求。

7.测试

7.1.测试策略

基于星火大模型 API 的法律诉状书生成系统的测试策略是确保系统的准确性、稳定性和可用性。

准确性：准确性关注系统是否能够准确地实现其功能和需求，确保输出与预期相符。需求验证，与法律同学合作，确保需求的正确理解和实现。单元和功能测试，对各个组件和整体功能进行深入测试，确保逻辑正确，输出准确。数据验证，确保使用准确的法律条款、案例和模板，与实际法律规范保持一致。

稳定性：稳定性关注系统在各种条件下的行为表现，包括负载、网络波动等。故障恢复测试，故意中断服务，测试系统的恢复能力和数据一致性。监控和报警，实施实时监控，确保及时发现和响应潜在问题。

可用性：可用性关注系统是否易用、访问方便，以及是否满足各种使用场景。用户界面测试，通过人机交互原则，测试用户界面的直观性和易用性。无障碍测试，确保系统对所有用户的可访问性，包括残疾人士。多平台兼容性测试，在不同设备和浏览器上测试系统，确保广泛的可访问性。可维护性和扩展性测试，通过代码审查和架构评估，确保系统易于维护和扩展。

7.2.测试用例

7.2.1.测试用例 1

原告 1：我叫李大勇，性别男，于 2010 年 12 月 22 日出生，汉族。我住在山东省

济南市历下区，手机号码 13712345678，身份证号是：120109201012220803，我的委托诉讼代理人人是山东东方律师事务所的律师郑建国，我的法定代理人是贾振扬。

原告 2：海翼人寿保险股份有限公司山东分公司济南中心支公司。住所地在山东省济南市绿地大道与长清路交叉口东方广场商务楼第 8 号。该公司的法定代表人是该公司的经理王立伟。（联系电话 12345678901）。统一社会信用代码：52719384621。委托诉讼代理人人是山东西方律师事务所的律师贾建国。

被告 1：她叫赵美丽，性别女，于 1989 年 3 月 8 日出生，汉族。住在湖北省武汉市武昌区，手机号码 15987654321，身份证号是：120109198903080323，她没有委托诉讼代理人，也没有法定代理人。

被告 2：天盛财产保险股份有限公司湖北分公司武汉中心支公司。住所地在湖北省武汉市解放大道与和平路交叉口新华大厦 1 号。该公司的法定代表人是该公司的经理李刚。（联系电话 23456789012）。统一社会信用代码：63827194035。委托诉讼代理人人是湖北朝阳律师事务所的律师冯国。

诉求：我的要求是：第一，被告需承担全部责任并赔偿我的车辆修理费用共计 20,000 元；第二，被告需赔偿我的医疗费和误工费共计 10,000 元；第三，本次诉讼的费用要由被告来承担。

事实经过：我与被告均为驾驶员，2019 年 11 月 12 日下午 5 点 30 分，当我驾驶车辆沿主路行驶至某十字路口时，被告驾驶的车辆突然从侧路闯入，未能及时刹车与我车相撞，造成我车严重受损并导致我受伤。具体说，我当时行驶在正常速度并且已在绿灯下通过，被告却未按规定在红灯下停车，直接闯入了我的行车道。被告随后承认了他的过错并答应赔偿，但之后一直未兑现。我多次与被告协商，未果，所以我决定起诉他。

证据：我有现场的监控录像，可以清楚地看到被告的过错行为；我还有与被告当天的通话记录和后续的短信聊天记录；另外，我还保留了医院的诊断报告和车辆的修理费用单据。

法院：湖北省武汉市武昌区人民法院

7.2.2.测试用例 2

原告 1：我叫李大勇，性别男，于 1983 年 12 月 22 日出生，汉族。身份证号是 320404198312222819。我住在历下区，手机号码 13712345678，

被告 1：她叫赵美丽，性别女，于 1989 年 3 月 8 日出生，汉族。住在武昌区，手机号码 15987654321，身份证号是 320132198903082322。她的委托诉讼代理人人是湖北金信律师事务所的律师王芳。

诉求：我的要求是：第一，婚后共同财产分割，被告应支付给我 20,000 元；第二，被告应支付赡养费 5,000 元；第三，被告应支付孩子的抚养费 10,000 元；第四，这次诉讼的费用要由被告来承担。

事实经过：我与被告结婚于 2010 年 5 月 1 日，育有一子。婚后，被告多次对我进行粗暴对待，并且经常外出夜不归宿。尽管我多次劝说，被告的行为却一直没有改变。在 2018 年 3 月 29 日，我发现被告与他人有婚外情。尽管我努力挽救婚姻，被告还是一意孤行，继续与第三人保持不正当关系。鉴于被告长期的不负责任行为和对家庭的破坏，我已无法继续与被告共同生活，所以我决定起诉离婚。

证据：我有被告的暴力行为的医院诊断记录和报警记录；我还有被告与第三人通话、短信及社交媒体聊天记录；另外，我还保留了我们的结婚证、孩子的出生证明和婚后共同财产的购买凭证等。

法院：武昌区人民法院

我们选取了两个测试用例，用来测试本系统是否能正常运行，其中包含多原告被告，多个人或公司，涵盖机动车交通事故纠纷和离婚纠纷，选取样例丰富多样，系统稳定性强。

7.3.测试结果

我们的聊天对话**最新的消息在最上面**，因此下面图像中的信息列表均是从下往上阅读。

下图给出了例 1 的部分截图，如图中所示，我们输入了一个人的原告信息，系统自动提取到了相关信息，用表格的形式输出了内容，同时输出量对应的语音。



下面一张图给出了事故发生的主要信息，大模型会自动识别出具体的案由，并且提供

相关的诉讼申请模板。用户可以根据给出的提示，结合自身情况，更为准确地描述自身案件的诉讼请求。



用例 2 部分截图：

1) 生成原告和被告信息结果：

好的，我已经知道第1个原告的信息了，请问你是否需要继续添加原告信息？	
描述	内容
姓名	李大勇
性别	男
出生日期	1983年12月22日
民族	汉族
住址	山东省济南市历下区
联系方式	13712345678
身份证号	320404198312222819
法定代理人	无
委托诉讼代理人	无

好的，我已经知道第1个被告的信息了，信息如下。请问你是否需要继续添加被告信息？	
描述	内容
姓名	赵美丽
性别	女
出生日期	1989年3月8日
民族	汉族
住址	湖北省武汉市武昌区
联系方式	15987654321
身份证号	320132198903082322
法定代理人	无
委托诉讼代理人_姓名	王芳
委托诉讼代理人_事务所	湖北金信律师事务所

7.4.测试结果分析

模型流畅性指的是本界面生成回答是否自然、连贯。由于该应用对用户操作流程要求比较严格，因而有时回答可能会显得啰嗦或者不够精确。

语音输入不能手动控制输入截止，使用起来不是很方便，但是在安静并且表达流畅的语境下，能够节省较多的输入时间。同时，输入结果可以修改，能够弥补因语音识别错误导致的问题。

测试中的错误可以分为两类：**识别错误**和**运行逻辑错误**。识别错误是指星火模型错误理解了问题或者提供了不准确的答案。运行逻辑错误则是当整个程序因为用户回答不合适导致的逻辑问题。在识别错误方面，星火模型可能会偶尔存在问题，因为模型的训练数据和对特定问题的了解是有限的。在逻辑运行错误方面，由于时间紧，任务重，团队磨合还不久，存在了一些上下文不匹配的情况，后期可以考虑使用更加科学的大模型开发方法。

8.思考与反思

8.1.用户交互体验设计

作为对话系统的设计者，用户交互体验是至关重要的考虑因素之一。为了提供良好的用户交互体验，我们需要关注语言风格、响应速度、用户反馈、操作简便性以及数据安全等方面。首先，语言风格应该亲切友好，尽量避免过于专业或复杂的术语，以确保用户理解和接受度。其次，需要及时、准确地回应用户，确保系统的实时性和效率。此外，交互过程应该简单明了。用户界面应该设计得直观易用，避免复杂的操作流程和冗长的提示信息，从而减少用户的学习成本和认知负担。最后积极倾听用户的意见和建议，保护用户的数据和隐私安全，以建立用户对系统的信任和依赖。

8.2.大模型在特定任务下的灵活性和可控性

灵活性是利用大模型完成更多更有创造力的任务，而可控性确保用户在一定引导下可以完成特定的任务。可控性在引导用户完成复杂流程任务情景下格外有效，比如法律文书生成任务。我们需要在大模型的开发和使用过程中，平衡灵活性和可控性。这可以通过设计合适的引导策略、设置任务限制和指导用户操作等方式来实现。同时，也需要不断反思和改进，以提高大模型在特定任务下的性能和可用性，从而更好地满足用户的需求。

8.3.Langchain 技术的应用

本次应用其实没有采用工程化的技术，LangChain 技术是一种基于语言模型构建应用程序的框架，提供了一系列工具和组件，简化了使用大模型技术的过程。下次如果构建基于语言大模型的应用，一定要采用现在很流行的 Langchain 技术流程。

9.附录

9.1.相关文档

1. 配置文档
2. 测试用例
3. prompt 设计文档
4. 生成起诉文书