**需求规格说明书**

**1. 引言**

**1.1 编写目的**

本需求规格说明书旨在详细阐述文言文处理工具的需求，为项目开发团队提供清晰、明确的需求指导，确保开发过程顺利进行。

**1.2 背景**

随着中华文化的复兴，文言文作为中国传统文化的载体，其研究与应用日益受到重视。然而，文言文复杂的语法结构和独特的表达方式使得一般用户难以深入理解。为此，开发一款能够对文言文进行智能处理的工具显得尤为重要。

**1.3 定义和缩略词**

* **文言文**：古代汉语的一种书面语形式。
* **分词**：将连续的文本切分成有意义的词汇单元。
* **词性标注**：为文本中的每个词标注其语法属性。
* **断句**：在文本中插入适当的标点符号，以标识句子的边界。
* **标点**：为文本添加标点符号，使其更具可读性。

**2. 总体描述**

**2.1 产品视角**

文言文处理工具是一款面向教育和学术研究领域的Web应用，旨在帮助用户更高效地理解和分析文言文。

**2.2 产品功能**

* 文言文分词
* 文言文词性标注
* 文言文断句
* 文言文标点

**2.3 用户特征**

* 教育工作者
* 学术研究者
* 文言文爱好者
* 学生

**2.4 约束条件**

* 系统应能在主流浏览器上运行，如Chrome、Firefox、Safari等。
* 系统后端应能在标准的Python环境中部署。

**3. 具体需求**

**3.1 功能需求**

**3.1.1 文言文分词**

* **输入**：用户输入的文言文文本。
* **处理**：调用后端分词模型进行文本分词。
* **输出**：分词结果，以列表形式展示。

**3.1.2 文言文词性标注**

* **输入**：用户输入的文言文文本。
* **处理**：调用后端词性标注模型进行词性标注。
* **输出**：词性标注结果，以列表形式展示。

**3.1.3 文言文断句**

* **输入**：用户输入的文言文文本。
* **处理**：调用后端断句模型进行断句处理。
* **输出**：断句结果，以列表形式展示。

**3.1.4 文言文标点**

* **输入**：用户输入的文言文文本。
* **处理**：调用后端标点模型进行标点添加。
* **输出**：添加标点后的文言文文本。

**3.2 界面需求**

**3.2.1 主界面**

* 提供文本输入框，供用户输入文言文文本。
* 提供四个功能按钮：分词、词性标注、断句、标点。
* 显示处理结果区域。

**3.2.2 结果展示**

* 以清晰、易读的格式展示处理结果。
* 提供原始文本和处理后文本的对比视图。

**3.3 性能需求**

* 分词、词性标注、断句、标点功能应在用户可接受的响应时间内完成。
* 系统能够处理不同长度的文本，而不会出现明显的性能下降。

**3.4 安全需求**

* 对用户输入进行验证，防止SQL注入、XSS攻击等。
* 确保用户数据的安全性，不泄露用户隐私。

**3.5 系统约束**

* 系统应能在常见的操作系统上运行，如Windows、Linux、macOS等。
* 系统前端应遵循W3C标准，确保良好的兼容性。

**4. 数据描述**

**4.1 数据对象**

* 文言文文本
* 分词结果
* 词性标注结果
* 断句结果
* 标点结果

**4.2 数据处理**

* 文本输入：用户通过前端界面输入文言文文本。
* 结果输出：后端处理完成后，将结果返回给前端界面展示。

**5. 附录**

**5.1 技术规范**

* 后端开发语言：Python
* 前端开发技术：HTML、CSS、JavaScript
* 数据传输格式：JSON

**5.2 参考文献**

* 甲言开源大模型文档
* Flask官方文档
* Bootstrap官方文档

**5.3 修订历史**

* 初始版本：1.0
* 修订日期：2024-10-27
* 修订人：冯森权