<UniGPT v2>

系统测试报告

版本 <1.0>

修订历史记录

|  |  |  |  |
| --- | --- | --- | --- |
| **日期** | **版本** | **说明** | **作者** |
| <09/08/2024> | <1.0> | 填写系统用例报告 | 陈启炜 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

目录

1. 简介 4

1.1 目的 4

1.2 范围 4

1.3 定义、首字母缩写词和缩略语 4

1.4 参考资料 4

1.5 概述 4

2. 测试概要 4

3. 测试环境 4

4. 测试结果及分析 4

4.1 需求覆盖率及缺陷分布 4

4.2 缺陷严重程度 5

5. 测试结论与建议 5

系统测试报告

# 简介

## 目的

## 本测试报告旨在验证UniGPT平台的功能和性能，确保其满足用户的需求，并识别可能存在的缺陷。

## 范围

该测试报告涵盖UniGPT项目的所有功能模块，包括AI提示词生成、插件交流平台等。文档影响的其他事项包括相关的用户体验和系统集成。

## 定义、首字母缩写词和缩略语

* **LLM**：大型语言模型（Large Language Models）
* **Prompt:** 指提示，即向模型提供输入以引导其生成特定输出的文本或指令，用于描述用户想要从模型获取的信息、回答、文本等内容。Prompt 的目的是引导模型产生所需的回应，以便更好地控制生成的输出。
* **Prompt Engineering**：提示工程，指优化输入LLM的提示以获得更好的输出结果的过程
* **API**：应用程序编程接口（Application Programming Interface）
* **GPT-4**: 生成型预训练变换模型 4（Generative Pre-trained Transformer 4）是由OpenAI公司开发并于2023年3月14日发布的自回归语言模型

## 参考资料

《<UniGPT v2>软件需求规约》 版本 <1.2>

《UniGPT System Test Case Form UniGPT v2系统测试用例》版本<1.1>

## 概述

* 本报告包含测试的目标、范围、方法、环境、结果及分析。文档结构分为测试概要、测试环境、测试结果及需求覆盖率分析等部分。

# 测试概要

* 测试于2024年8月1日至8月5日在线上连接服务器进行。参与人员包括项目开发团队和测试团队（陈启炜，韦东良，黄峻涛，叶懿芯）。测试方法采用黑盒测试和白盒测试相结合，测试内容涵盖功能测试，性能测试，功能测试，易用性测试，可靠性测试，安全性测试，兼容性测试，界面测试，压力测试和部署测试。
* 。

# 测试环境

通过本地主机（12th Gen Intel(R) Core(TM) i5-12500H，16Gi内存 DDR4. 0000:02:00.0 Ethernet controller: Realtek Semiconductor Co., Ltd. RTL8111/8168/8411 PCI Express Gigabit Ethernet Controller (rev 15) ) 使用JMeter, cypress等工具进行测试。

# 测试结果及分析

|  |  |  |  |
| --- | --- | --- | --- |
| 类别 | 用例数 | 通过数 | 通过率 |
| 功能测试 | 40 | 40 | 100.00% |
| 性能测试 | 37 | 37 | 100.00% |
| 易用性测试 | 1 | 1 | 100.00% |
| 可靠性测试 | 1 | 1 | 100.00% |
| 安全性测试 | 2 | 2 | 100.00% |
| 兼容性测试 | 1 | 1 | 100.00% |
| 界面测试 | 14 | 14 | 100.00% |
| 压力测试 | 7 | 7 | 100.00% |
| 部署测试 | 1 | 1 | 100.00% |
| 合计 | 104 | 104 | 100.00% |

## 需求覆盖率及缺陷分布

**表1 需求覆盖率及缺陷分布**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **测试内容** | | **测试**  **用例数** | **需求**  **覆盖率** | **缺陷数** | **缺陷率** | **备注** |
| **功能项** | 用户登录 | 2 | 100% | 0 | 0 |  |
| 创建插件 | 1 | 100% | 0 | 0 |  |
| 将插件发布到应用市场 | 1 | 100% | 0 | 0 |  |
| 大文本输入框 | 5 | 100% | 1 | 0.2 | 微小错误 |
| 输入字符长度验证 | 1 | 100% | 0 | 0 |  |
| 必填项验证 | 1 | 100% | 0 | 0 |  |
| 快捷键检查 | 1 | 100% | 0 | 0 |  |
| 进行注销操作 | 1 | 100% | 0 | 0 |  |
| 错误页面测试 | 1 | 100% | 0 | 0 |  |
| 多列表选择框 | 1 | 100% | 0 | 0 |  |
| 格式、规则输入验证 | 1 | 100% | 0 | 0 |  |
| 单选框、多选框的输入验证 | 1 | 100% | 0 | 0 |  |
| 系统错误定位的输入验证 | 1 | 100% | 0 | 0 |  |
| 页面链接检查 | 1 | 100% | 0 | 0 |  |
| 相关性检查 | 1 | 100% | 0 | 0 |  |
| 检查按钮的功能是否正确 | 1 | 100% | 0 | 0 |  |
| 重复提交表单 | 1 | 100% | 0 | 0 |  |
| 多次IE回退 | 1 | 100% | 0 | 0 |  |
| 上传下载文件检查 | 1 | 100% | 0 | 0 |  |
| 进行注销操作 | 1 | 100% | 0 | 0 |  |
| 删除操作 | 1 | 100% | 0 | 0 |  |
| 增加操作 | 1 | 100% | 0 | 0 |  |
| 修改操作 | 1 | 100% | 0 | 0 |  |
| 查询操作 | 1 | 100% | 0 | 0 |  |
| 翻页操作 | 1 | 100% | 0 | 0 |  |
| 整体风格一致性验证 | 1 | 100% | 0 | 0 |  |
| 窗体一致性验证 | 1 | 100% | 0 | 0 |  |
| 按钮一致性验证 | 1 | 100% | 0 | 0 |  |
| 输入框一致性验证 | 1 | 100% | 0 | 0 |  |
| 访问方法使用验证 | 1 | 100% | 0 | 0 |  |
| 页面链接验证 | 1 | 100% | 0 | 0 |  |
| Script脚本验证 | 1 | 100% | 0 | 0 |  |
| 与机器人交互并查看对话历史 | 1 | 100% | 0 | 0 |  |
| 创建机器人 | 1 | 100% | 0 | 0 |  |
| 将机器人发布到应用市场 | 1 | 100% | 0 | 0 |  |
| 登录 | 1 | 100% | 0 | 0 |  |
| 系统管理 | 1 | 100% | 0 | 0 |  |
| 创建插件 | 1 | 100% | 0 | 0 |  |
| 将插件发布到应用市场 | 1 | 100% | 0 | 0 |  |
| 创建机器人 | 1 | 100% | 0 | 0 |  |
| **功能项小计** | 40 | 100% | 1 | 0.025 |  |
| **非功**  **能项** | 性能测试 | 37 | 100% | 1 | 0.027 |  |
| 易用性测试 | 37 | 100% | 0 | 0 |  |
| 可靠性测试 | 1 | 100% | 0 | 0 |  |
| 安全性测试 | 1 | 100% | 0 | 0 |  |
| 兼容性测试 | 2 | 100% | 0 | 0 |  |
| 界面测试 | 1 | 100% | 0 | 0 |  |
| 压力测试 | 14 | 100% | 0 | 0 |  |
| 部署测试 | 7 | 100% | 0 | 0 |  |
| 性能测试 | 1 | 100% | 0 | 0 |  |
|  | **非功能项小计** | 64 | 100% | 1 | 0.015 |  |
|  | **总计** | 104 | 100% | 2 | 0.019 |  |

## 缺陷严重程度

**表2 缺陷严重程度**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **严重程度** | **致命** | **严重** | **一般** | **细微** | **总数** |
| **缺陷个数** | 0 | 0 | 0 | 2 | 2 |
| **占缺陷百分比** | 0 | 0 | 0 | 0.019 | 0.019 |

# 测试结论与建议

测试过程中遇到的问题：

在性能测试中，我们发现测试结果的波动较大，导致难以精确衡量平台的性能。可能的原因包括测试环境的不稳定、网络波动或系统资源未充分优化。此问题需要进一步调查，以确保一致性和可靠性。

测试结果分析

尽管如此，测试总体完成程度较好。功能模块基本实现了预期需求，用户界面友好，交互流畅。但仍存在一些细微缺陷，如某些边缘用例在复杂输入条件下的响应不够理想。

结论

总体而言，UniGPT平台已初步具备交付条件，但仍需注意以下几点：

1. 功能细节: 需关注功能实现的细节，确保在所有可能的用户交互中表现一致。

2. 接口性能优化: 某些API接口在高负载下表现不佳，建议进行性能分析和优化，可能需要改进数据库查询或增加缓存机制。

3. 环境稳定性 : 为提高性能测试的准确性，建议在更稳定的环境中进行测试，或使用模拟工具来减少外部变量的影响。

建议措施

1. 性能监控和优化: 引入性能监控工具，定期分析性能瓶颈，优化代码逻辑和数据库结构。

2. 测试环境改进: 创建标准化的测试环境，减少环境变动对测试结果的影响。

3. 持续改进和反馈: 建立反馈机制，持续收集用户反馈和异常报告，通过迭代更新不断提升平台质量。

通过以上措施，UniGPT平台将更好地满足用户需求，并提升整体用户体验。