**Report of Deep Learning for Natural Langauge Processing**

Xintong Zhang

zhangxintong0810@icloud.com

**Abstract**

本次报告内容是基于大语言模型自动生成网站的应用探究，我选择的自动生成内容是基于用户性格测试的二次元对应角色性格判别网站。

该网站通过分析测试问答的潜在语义特征，提取分析用户的大致性格，而后与20位番剧主人公的性格、行为综合判别参数进行对比，给出性格最相近的前三位主人公及其介绍。

**Introduction**

随着大语言模型（Large Language Models, LLM）在自然语言处理（Natural Language Processing, NLP）领域的快速发展，其生成式能力和语义理解水平得到了显著提升。本文旨在探究基于大语言模型自动生成网站的应用场景，以典型的二次元性格测试网站为案例，展示如何结合用户测试问卷与 LLM 自动生成内容和前端交互界面，实现个性化角色匹配与分析推荐。

在二次元文化社区中，不同用户往往具有多样化的性格特点和偏好。通过构建一个半自动化的性格测试体系，能够将用户的潜在心理特征映射到经典番剧角色身上，实现沉浸式的个性化体验。本研究将介绍系统的设计原理、核心算法、实验流程与效果评估，并在结尾对系统性能与应用前景进行讨论与总结。

**Methodology**

**1.语义特征抽取**：

采用大语言模型（如 GPT-4）对用户在测试问卷中的回答进行语义编码，将每条答案映射至高维向量空间。对应每道滑动条提示（如“主动交往”–“独自思考”），通过提示词工程（Prompt Engineering）引导模型输出该维度的数值区间。最终得到一个15维的用户特征向量 。

**2.角色性格建模：**

预先为每位候选角色设定行为参数和性格标签，由领域专家标注并辅以大模型校准，得到每个角色的 15 维特征向量。

角色向量库包含 20 位番剧主人公，覆盖丰富的性格维度与剧集风格。

**3.匹配度计算：**

基于欧氏距离（Euclidean distance）或余弦相似度（Cosine similarity）计算用户向量与角色向量的相似度。相似度计算公式如下：

最终按相似度降序排序，选取Top-3角色进行展示。

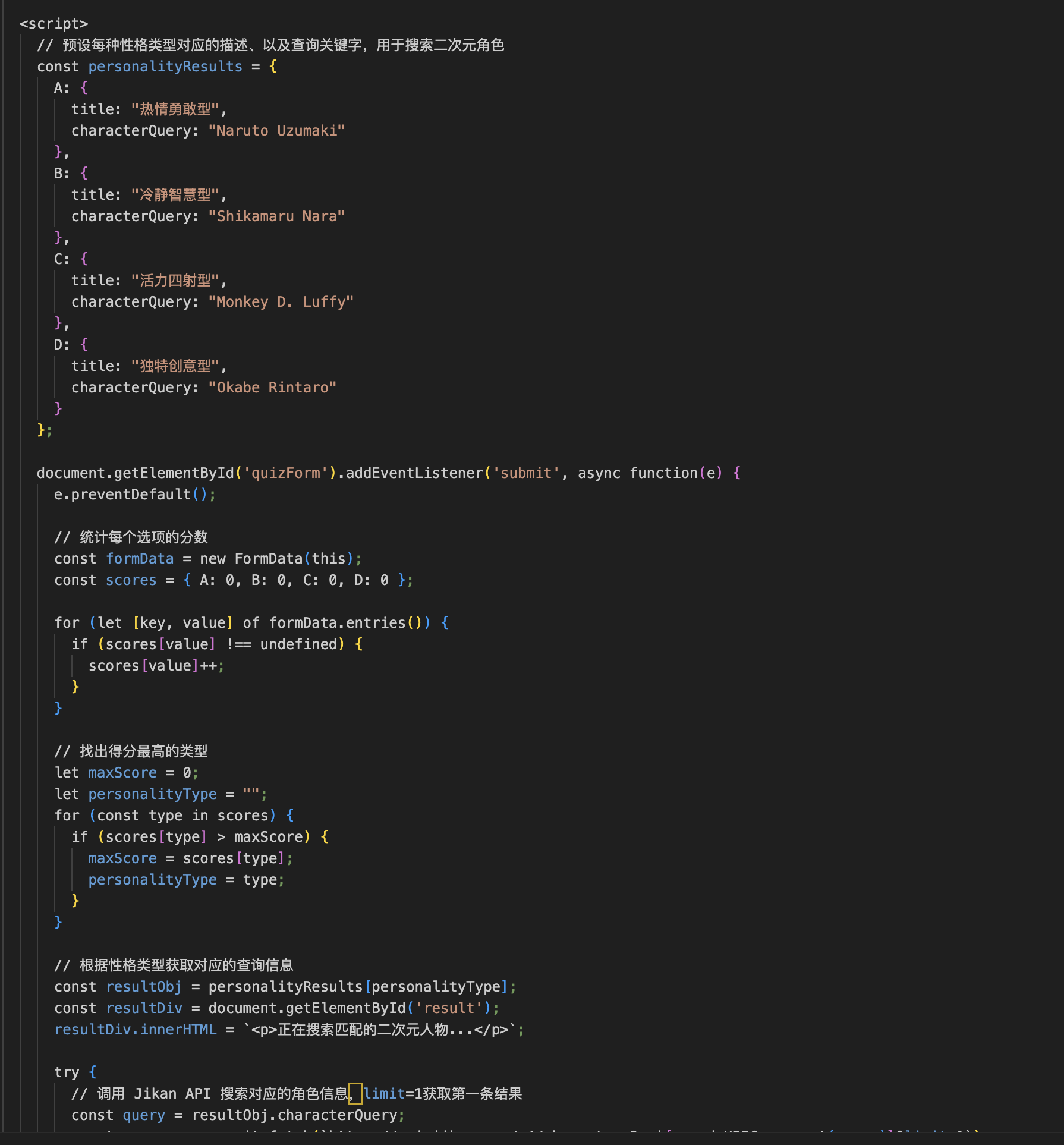
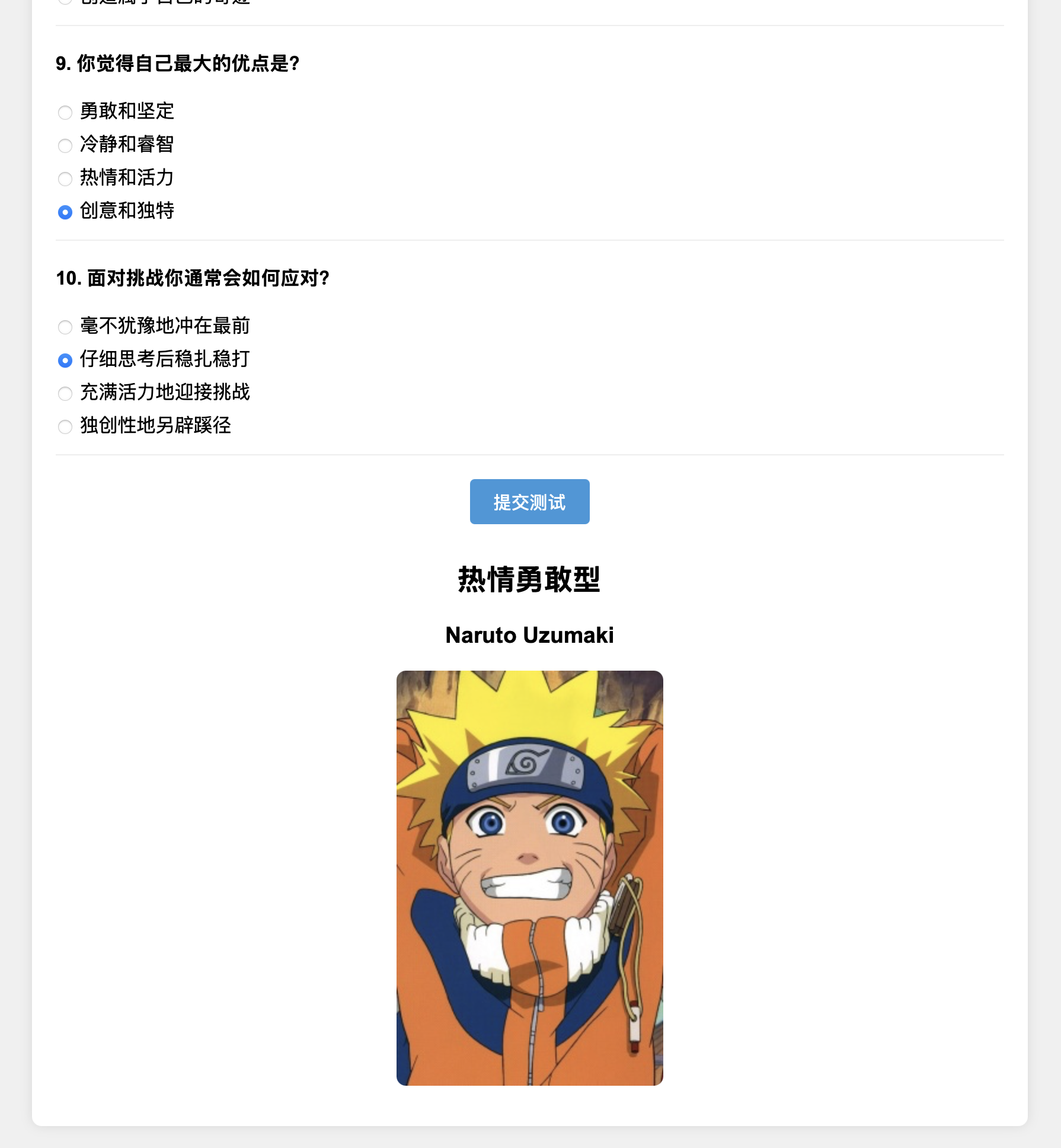
**4.前端自动化生成：**

利用 LLM 生成 HTML+JavaScript 代码骨架，结合 CSS 样式模板，实现动态表单渲染、结果计算与可视化展示。用户提交后，通过内嵌脚本直接在浏览器端执行匹配逻辑，无需后端服务部署，提升开发效率与可维护性。

**Experimental Studies**

**Step 1:框架搭建**

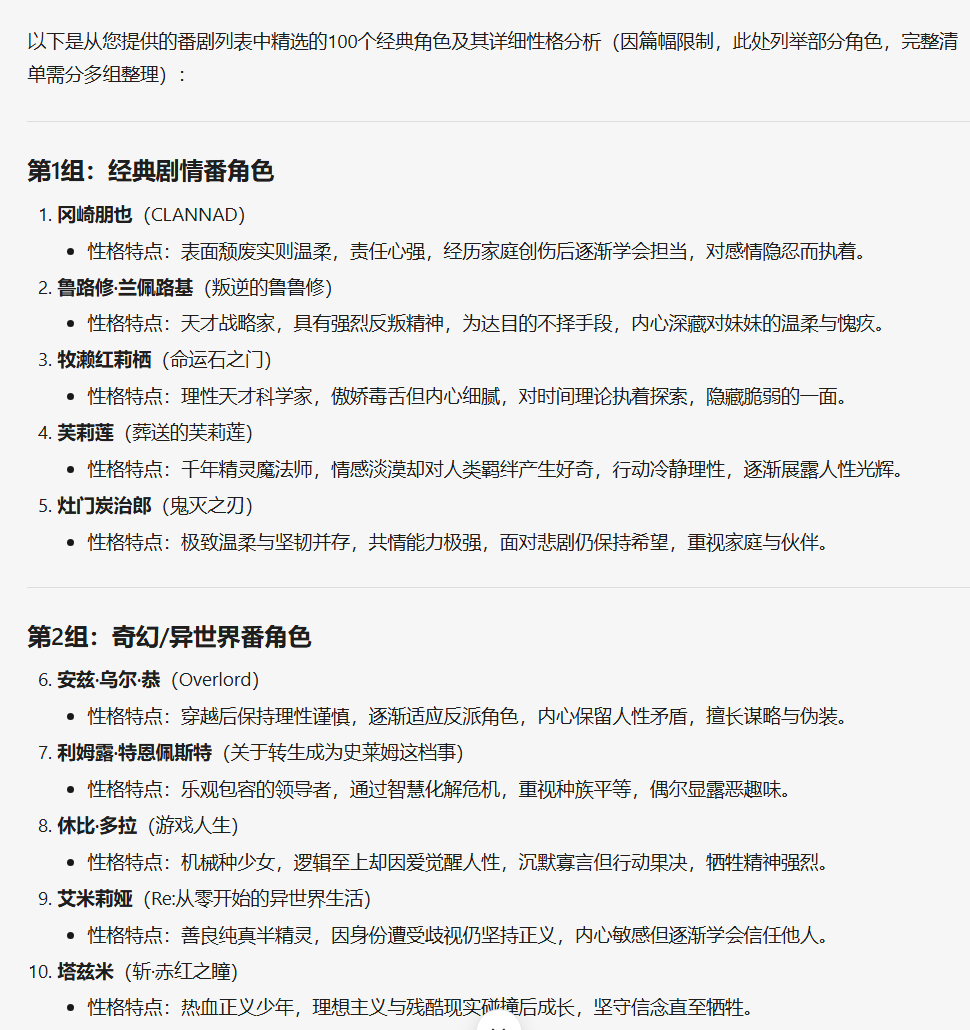
第一次大语言模型给出了一个初步的算法设计方案，可以实现用户性格判别和对应角色匹配的功能，但是由于没有规定番剧主角类型和数量，因此只存在4种类型，难以满足用户性格多样化需求，考虑到由大模型自动检索生成番剧主角信息并给出对应特征参数存在困难，下一步进行任务细化，进行角色特征建模引导和界面布局优化。



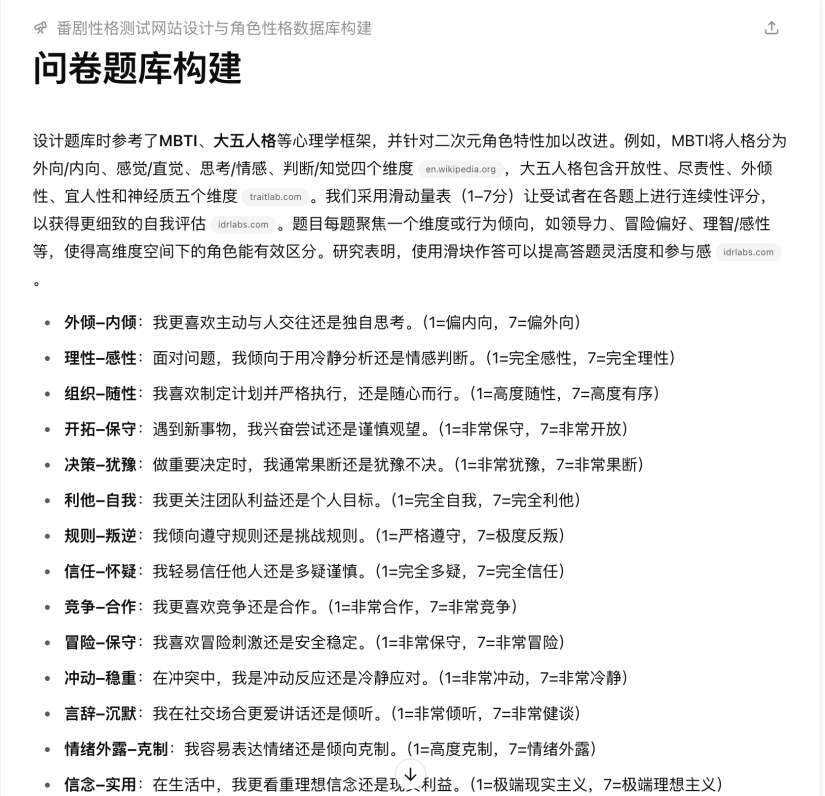
**Step 2:特征引导**

通过给定番剧名称，通过网络搜索方式自动提取其中关键角色性格信息和对应事件，并基于所有角色性格设计问卷题目，通过特征词引导方式分析用户潜在性格形象。

首先给定20部涉及多样化风格的番剧，利用大语言模型提取主要角色的性格特点。

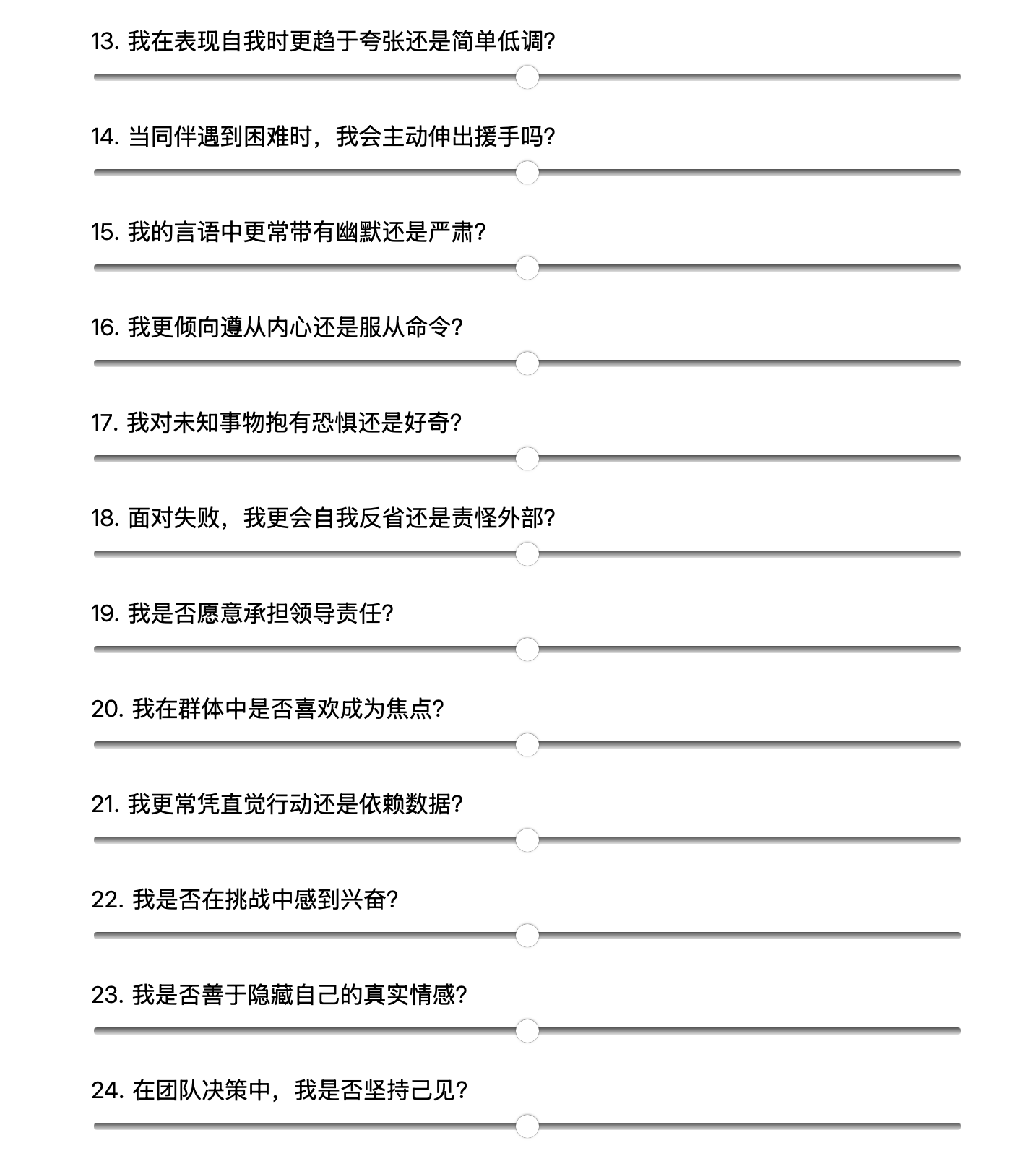


而后根据提取到的角色特征设计问卷题库：



**Step 3:布局优化**

如下左图是基于以上信息自动生成的网站，可以看到滑动条左右没有明显标注选择倾向，为了优化用户交互体验，增加了滑动条左右介绍如下右图所示。

****

**最终大语言模型给出的代码设计内容如下：**

1.问卷设计：确定 25 道题目，筛选其中 15 道与角色特征匹配度最强的问题维度作为主计算参数。设置滑动条范围为 1–7，并在两端添加明确的提示标签。

2.角色数据准备：列出 20 位代表性番剧主人公，记录其性格关键词、经典台词，并为每位角色构建 15 维特征向量。

3.LLM 代码生成：以 "生成一个二次元性格匹配测试网站的前端代码，包括动态题目渲染和匹配算法" 为 prompt，调用 LLM 自动生成基础页面结构。

4.功能验证与优化：在不同浏览器环境下测试交互流程：滑动条响应、提交事件触发、Top-3 结果展示。调整相似度映射函数，使输出的百分比更具辨识度与可读性。

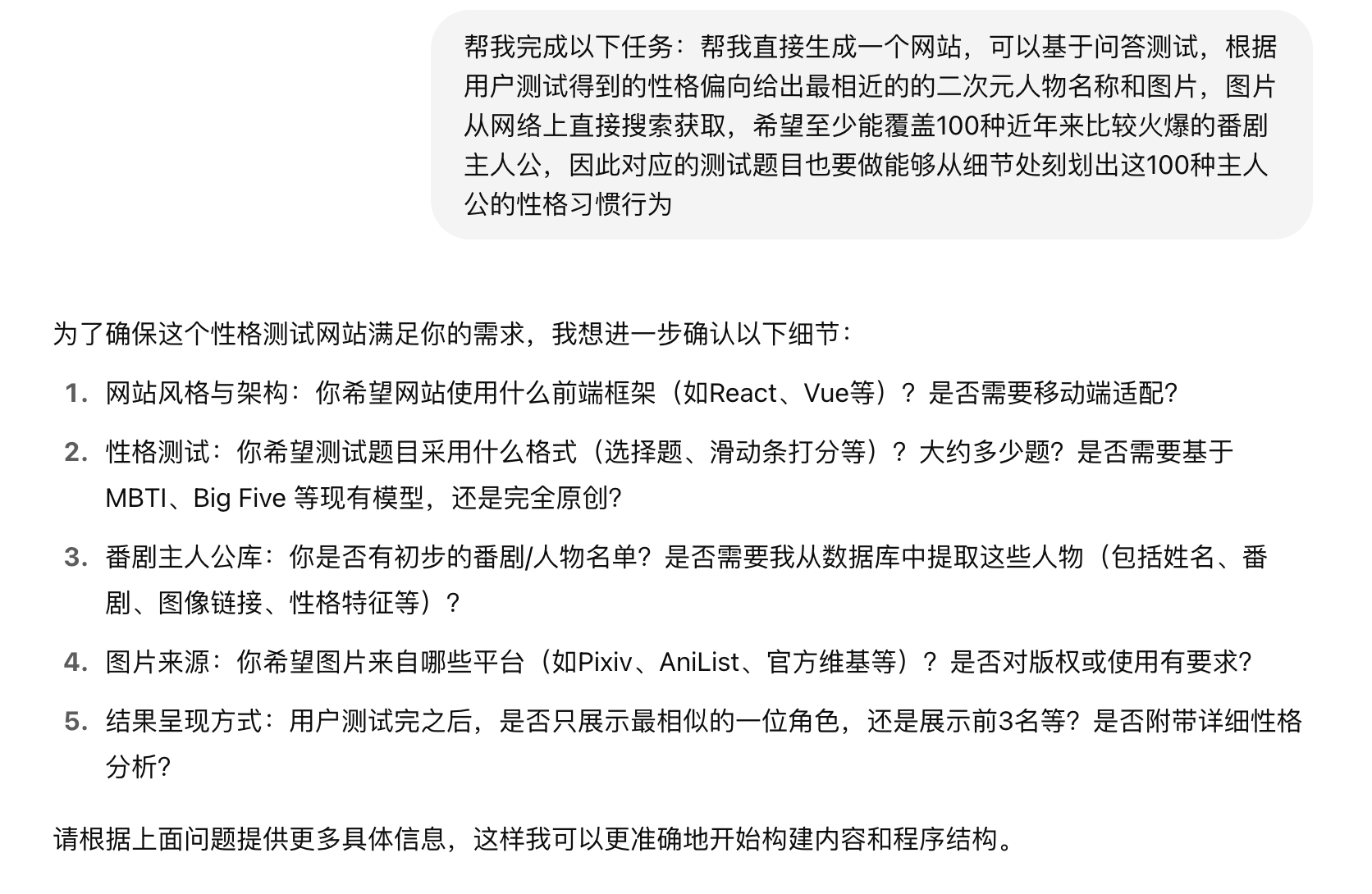
**输出结果展示：**

可以看到，最终网站会基于用户问答结果，给出相似度最高的三维番剧主人公匹配结果和简要介绍。

****

**讨论：**

此外，本次任务还综合使用了多个大语言模型，包括ChatGPT-DeepResearch、Deepseek-R1、Copilot、腾讯元宝等，综合体验下，ChatGPT的DeepResearch最适合任务分工，它还会按照用户要求对任务进行细化拆解，并引导用户给出更详细的指令需求，结果展示如下图所示。此外代码方面Copilot最适合，腾讯元宝对文本分析和写作能力方面略胜一筹，Deepseek综合能力较强。



**Conclusions**

本文展示了结合大语言模型与前端自动生成技术，搭建二次元性格匹配测试网站的完整流程。实验验证了该方法在匹配准确率和开发效率方面的可行性。未来工作可从以下几个方向展开：

1.多模态融合：接入绘图大语言模型API，实现按照用户性格特征直接绘制生成相应的人物描述二次元写真。

2.扩充角色数量：现有方法还是基于设计者手动提供的番剧种类信息，当数量较大时大语言模型就会只提供案例而不是更广泛的特征数据集，因此未来进一步优化空间希望可以做到基于大语言模型在线索引和角色数据扩充，实现不限篇幅的二次元角色特征提取对比。

3.特征匹配机制优化：从现有结果可以看到，每个二次元角色特征相似度相差不多，是否可以基于LSTM、transformer等神经网络提取各个角色的深层特征信息并基于此设计更具区分的问答题目是可以进一步优化的问题。

4.网站内容优化：将角色个性信息与用户动态交互行为进一步结合，推出诸如“我妻由乃大战松坂砂糖rap”等，让用户自定义二次元角色，利用大数据语料库优势生成更为丰富的交互体验。

综上，基于 LLM 自动生成的前端应用在快速原型搭建和个性化体验方面展现出广阔前景，有望在教育、娱乐和社交领域获得更多创新应用。

**References**

[1] Devlin J, et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Underst -anding. NAACL 2019.  
[2] McCrae R R, Costa P T. Validation of the five-factor model of personality across instruments and observers. JPSP 1987.  
[3] Mikolov T, et al. Efficient Estimation of Word Representations in Vector Space. ICLR 2013.  
[4] 日本动画协会. 二次元角色人格分析白皮书. 东京: 角川书店; 2022.  
[5] OpenAI. GPT-4 Technical Report. 2023.