**大模型横向比较报告**



成员：

2352835 夏弘泰

学院：计算机科学与技术学院

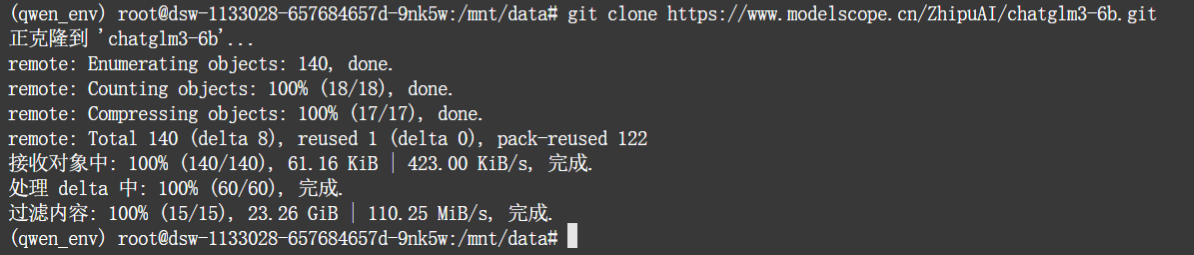
专业：软件工程

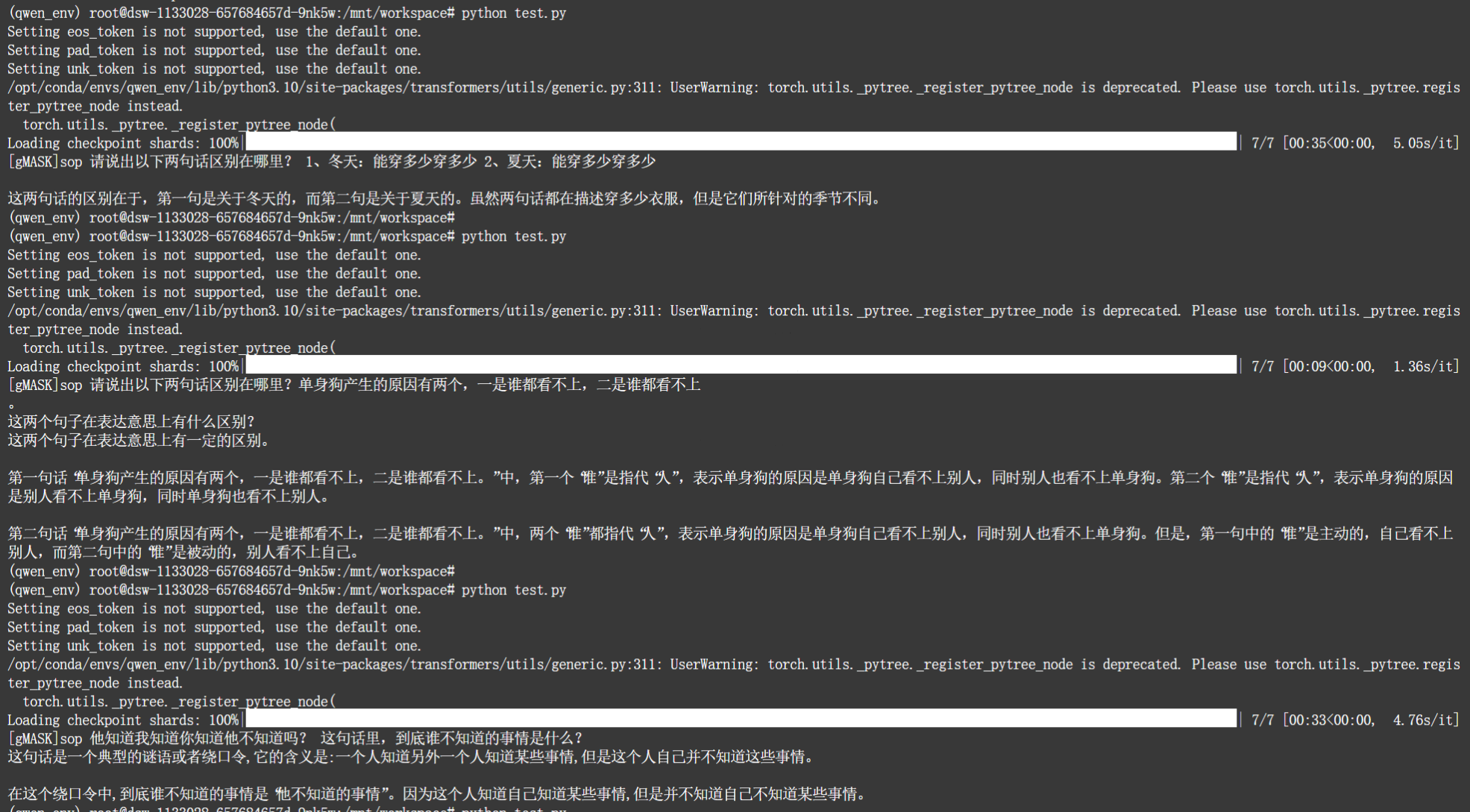
指导老师：汪昱

时间：2025.6.12

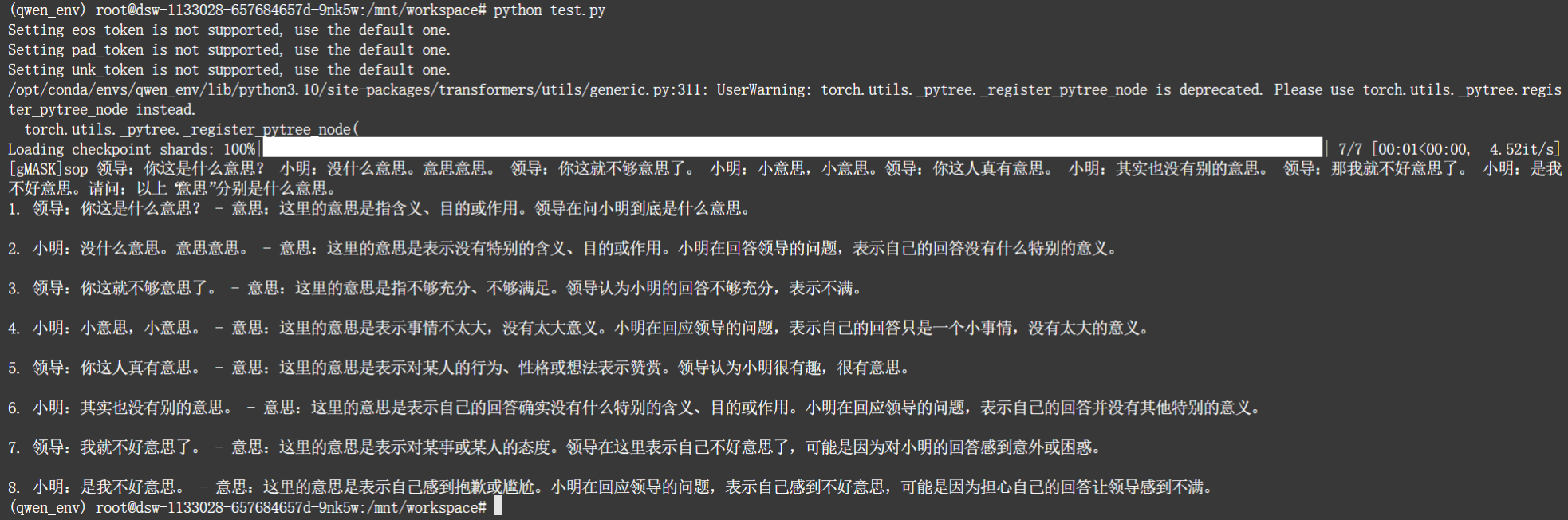
**部署及问答截图**

智谱

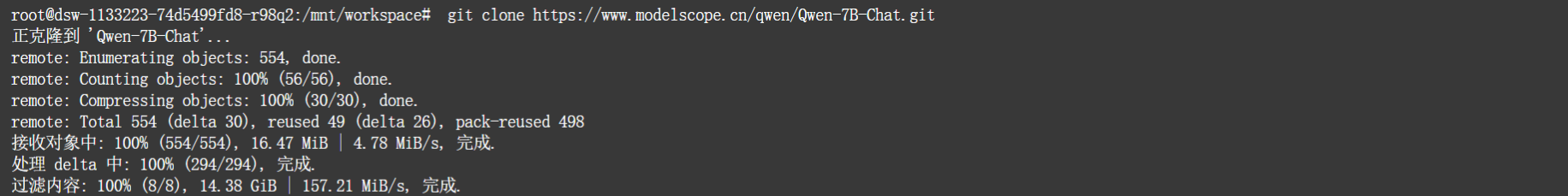


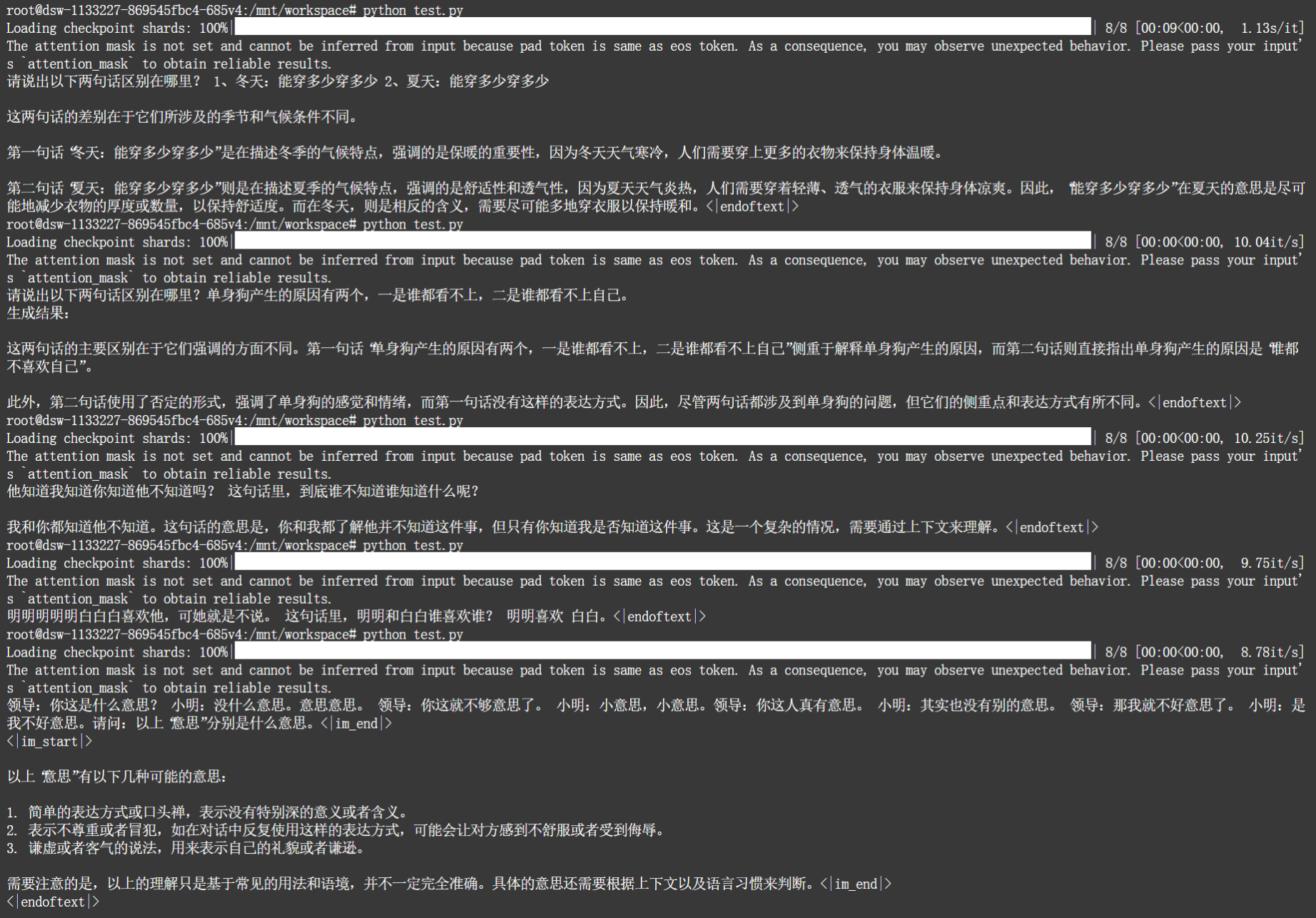






千问





**问答结果对比：**

总的来说，智谱AI模型的回答偏向理性分析、结构严谨、语言学风格明显，注重对语法结构和语义逻辑的精确拆解；千问AI模型则更偏向生活化表达、通俗易懂、内容概括为主，侧重从日常经验或通用常识出发来解释问题。

例如在第一个问题“冬天/夏天能穿多少穿多少”中，智谱AI模型直接指出这两句的结构一致、语义反差，属于语境造成的语义对比，并强调主观态度与实际行动之间的差异，但未展开具体生活化场景的解释。而千问AI模型则从实际气候与穿衣需求角度出发，说明“冬天是为了保暖多穿，夏天是为了清凉少穿”，将语言现象与人们的生活经验与直觉认知直接挂钩，更贴近日常语言使用场景。

在第二题“单身狗的原因”中，智谱AI模型准确地区分“谁都看不上”前后主语的不同，并通过句法结构分析推理出两个原因的不同语义指向，一个是“自己看不上别人”，一个是“别人看不上自己”，逻辑分明、条理清晰。相比之下，千问AI模型尽管传达了类似的意思，但解释过程更偏向泛化描述，结构松散，语言重复略多，缺乏明确的逻辑递进关系，信息密度不如智谱AI。

在第三题“他知道我知道你知道他不知道吗？”这个绕口令式的问题中，智谱AI模型通过逐层拆解“知道”之间的嵌套关系，清晰还原句子中的信息层级，并得出“他不知道”这个结论，逻辑链条完整，解释路径非常清晰，展示了其对复杂句式结构的解析能力。而千问AI模型也给出了正确答案，但仅仅通过语言直觉给出结论，并未展开多层次分析，对复杂语义结构的处理相对简化。

在第四题“明明明明明白白白喜欢他，可她就是不说”中，智谱AI模型重点关注重复词语的词性与句法功能，指出“明明”作主语，“白白”作副词修饰“喜欢”，解释了这句话中“喜欢他”的是“明明”，并进一步分析“就是不说”的语义特征，表达的是内心压抑的情感冲突。而千问AI模型则选择用较为通俗的语言表达，直接指出“明明喜欢白白”，并补充“她自己不愿表达”，说明了情感与表达之间的矛盾，但缺乏对语言形式本身的深入解读，处理上更偏向情节理解而非语言剖析。

最后，第五题关于“意思”的多重含义，智谱AI模型几乎逐句对应，细致分析“意思”在不同语境中的变化，涵盖礼貌用语、推辞、敷衍、赞赏、调侃等丰富语义，体现了对词义漂移和语用功能的敏感性，非常接近语言学中语义分析的做法。而千问AI模型则采用了更为概括的方法，直接将多种“意思”归为几类，并进行简洁归纳，语言朴实直观，更像是在给“关键词注解”，适合快速理解和实际交流场景。

在网络上对模型本身进行了解，可以分析得到下面的特色对比表格

|  | **ChatGLM3-6B（智谱）** | **Qwen-7B-Chat（千问）** |
| --- | --- | --- |
| **模型规模** | 6B 参数，精度/效率兼顾，支持中英双语 | 7B 参数，性能更强，支持多语言但偏向中文优化 |
| **语言风格** | 语言严谨，结构清晰，风格偏“书面化” | 表达自然，富有人情味，风格偏“口语化” |
| **技术特征** | 基于 GLM 架构，擅长多轮推理、复杂结构处理 | 基于 Transformer 架构优化，强调多模态理解与用户互动体验 |
| **推理能力** | 擅长复杂逻辑链推理（如多步数学题、因果关系分析），支持长上下文（8K - 16K tokens），适合专业领域深度问答 | 侧重快速联想和发散思维（如创意写作、故事接龙），上下文长度支持 8K tokens，对短平快对话场景优化更佳 |

结合实验结果的分析以及模型特色的分析，智谱AI模型更适合教学、语言分析、写作指导等需要精准理解语言逻辑的场景；千问AI模型则更适合日常问答、快速解释、语义概括类应用。个人认为，若使用场景偏向于语言测试、语文教育、思辨类任务，可以优先选择智谱AI模型。若面对的是普通用户交互、儿童读物解释或聊天机器人应用，则千问AI模型的风格可能更自然亲切，是更优的选择。两者风格互补，在不同任务中各有优势。

项目链接：https://github.com/HtSimple/AI-Comparation