

基于文心大模型的 AI 中文语伴开发模式对比研究 (A Comparative Study on Development Models of AI Chinese Language Partners Based on the ERNIE Large Language Model)

连维琛 (Lian, Weichen) 北京语言大学 (Beijing Language and Culture University) lwc19990514@gmail.com	郑明鉴 (Zheng, Mingjian) 北京语言大学 (Beijing Language and Culture University) Tengumorso@outlook.com	徐娟 (Xu, Juan) 北京语言大学 (Beijing Language and Culture University) xujuan@blcu.edu.cn
--	---	---

摘要: 语伴在语言学习中的重要性不言而喻,在生成式人工智能技术飞速发展的时代背景下,本研究尝试依托大语言模型构建 AI 中文语伴,选用百度文心一言大模型(ERNIE)作为基础大模型,通过编写提示词(Prompt)、集成本地中文教学资源等方式使通用大模型更适用于国际中文教育领域。本研究分别用低码态与代码态两种开发模式开发面向中文学习者的 AI 中文语伴,并对两种开发模式的开发过程进行对比,然后从基本功能、生成内容等方面对开发的 AI 中文语伴进行分析并给出择用建议。

Abstract: The importance of language partners in language learning is undeniable. In the context of the rapid development of generative artificial intelligence technology, this study attempts to construct an AI Chinese language partner based on large language models, utilizing the ERNIE model as the foundational model. By crafting prompts and integrating local Chinese teaching resources, the general model is tailored to better suit the field of international Chinese education. This research develops AI Chinese language partners for Chinese learners using both low-code and code-based development modes, comparing the development processes of these two approaches. Furthermore, the study analyzes the developed AI language partners in terms of basic functionalities and generated content, providing recommendations for their selection and use.

关键词: 大语言模型、AI 中文语伴、低码态、代码态、国际中文教育

Keywords: LLM, AI language partner, low-code, code-based, Chinese language education

1. 引言

语言伙伴 (Language Partner), 即语伴, 通常指与语言学习者一起学习语言的伙伴 (陈倩, 2020), 他们不仅可以是语言学习者, 也可以是练习语言的对象, 其作用主要体现在帮助二语学习者巩固语言知识、锻炼二语思维、提高交际能力 (周欣宜, 2024)。在国际中文教育领域, 一方面, 由于时间、空间、交流意愿、语言水平等多种因素, 在非洲、美洲等“非中文社区”, 中文语伴资源明显短缺, 为中文学习者匹配到合适的真人语伴并不容易; 另一方面, 专业语伴系统的开发较为复杂, 较难实现个性化, 且中文学习者直接使用传统的问答系统和通用生成式人工智能 (如 ChatGPT、文心一言 (ERNIE)) 并无法满足与真人语伴对话类似的效果, 效果不佳 (郑永红等, 2024), 其原因在于传统问答系统存在自然语言理解能力差、数据质量无法保证、维护成本高等问题 (闫悦等, 2023)。同时, 直接使用通用生成式人工智能则无法对生成内容的适需性、准确性进行有效控制, 无法满足专业领域的教育需求 (张春红等, 2023)。

基于大语言模型和智能体平台构建的智能体 (Agent), 有望能够担任语伴的角色, 为中文学习者创造“无人”环境下的中文交流机会, 填补“真人语伴”的空缺; 同时还可以借助大语言模型出色的自然语言理解和生成能力, 为中文学习者提供更适需、准确的语言学习服务。在国际中文教育领域, 曹钢 (2023) 和马瑞凌 (2024) 等人已在之前提出尝试采用“低码态” (Low-code) 或“代码态” (Code-based) 两种人机共创的开发模式建设各式新型的数字应用的想法, Trigo (2022) 等人于其他领域对两种模式下开发的应用进行了探索。据此, 本研究将实验中基于生成式人工智能构建中文语伴的开发模式划分为低码态和代码态两种, 并对两种开发模式下的 AI 中文语伴进行分析比较, 探讨以下两个问题:

- (1) 不同开发模式下 AI 中文语伴的具体构建过程为何?
- (2) 不同开发模式构建的 AI 中文语伴生成效果如何?

2. 相关概念界说及研究思路

2.1 低码态与代码态的界说

基于大语言模型的中文 AI 语伴开发可分为两种优势不同的模式: 低码态开发 (Low-code Development) 和代码态开发 (Code-based Development), 这两种开发模式适用于不同开发情况和需求。

第一, 低码态是一组数字技术工具平台, 基于图形化拖拽、参数化配置等更为高效的方式, 实现快速构建、数据编排、连接生态等 (Rokis & Kirikova, 2022), 旨在通过不用代码或少量代码 (如在组件中编写代码处理输入变量来生成返回值) 实现不同场景的构建, 通过图形工具简化了开发过程。低码态允许开发者对代码只

有基础了解甚至不需要了解, 开发者只需要根据任务需求使用平台提供的组件实现数字化转型中的场景应用创新。

第二, 与之相对应的则是代码态, 即传统的以编写程序来实现软件功能的开发模式, 例如通过代码对模型进行微调、界面设计、端口连接等, 具有高度的编码灵活性, 适用于各行业领域, 功能实现较灵活自由, 代码态开发要求开发者具备与代码相关的专业知识与丰富的编程背景, 通常由专业的技术人员来完成开发任务, 工作量较大、开发周期较长、成本较高。

综上, 从开发要求来看, 低码态开发模式允许代码能力不强甚至没有代码能力的中文教师能够通过设计提示词 (Prompt)、增加外部知识库等操作为中文学习者构建适需的 AI 中文语伴。代码态开发模式则允许通过编程对 AI 中文语伴进行更精细化、更具针对性的定制, 例如: 交流界面设计, 生成随机性、文本多样性等参数设置等。

2.2 智能体及其参数界说

智能体是指一种人工智能代理, 具备环境感知能力, 能够自主理解、决策和执行任务 (Wooldridge & Jennings, 1994)。它根据任务目标独立进行规划, 将复杂任务分解为多个步骤, 从而完成目标。

在构建 AI 中文语伴的过程中, 智能体的参数设定十分重要, 生成式人工智能的反馈会根据参数设定的不同而变化 (Mizumoto & Eguchi, 2023)。低码态和代码态下的参数配置各具特点, 为不同需求的用户需求提供了灵活的选项。在低码态的默认设定中, 参数简化并优化为直观的操作, 例如多样性调节 (Temperature) 和采样范围调节 (Sampling Diversity), 其中多样性参数控制生成文本的丰富性, 取值越大, 输出越具创意且多样; 采样范围则调节模型的随机性, 较高的数值带来更多变的回答, 而较低的数值则让输出趋于一致。

在代码态中, 参数配置更为精细, 为二语学习者提供了更广泛的个性化设置空间。例如设置惩罚 (Penalty) 用于减少生成内容中重复 token 的现象, 值越高, 重复的生成内容会被更强烈地抑制; 设置流式输出 (Streaming Output) 控制生成内容的输出速度, 让用户可以边生成边查看, 提升响应体验; 设定人设 (Persona) 让模型根据设定的个性或身份特征输出符合特定“人设”的内容, 以实现更个性化的互动。

2.3 网络爬虫技术界说

网络爬虫技术 (Web Scraping) 是一种用于自动化数据收集的技术手段 (盛景等, 2024), 广泛应用于获取互联网上的大规模信息。通过模拟用户浏览行为, 爬虫程序发送请求至服务器, 接收并解析返回的网页数据, 以提取特定内容。近年来, 爬虫技术开始帮助国际中文教育领域的研究者和教育工作者获取多源数据, 为中文学习资源库的建设、教学内容的更新以及学习者行为分析提供数据支持。

2.4 研究思路

第一, 本研究将基于研究目标设计 AI 中文语伴需具备的功能, 如社交功能、教学功能、辅助功能等, 根据应用功能对大模型进行选择并整理相关中文专业知识, 为后续应用开发做好准备。

第二, 运用两种开发模式进行具体的 AI 中文语伴开发实践, 其中低码态模式主要依据平台提供的选项, 借助功能描述、提示词设计、角色设定等方法, 完成 AI 中文语伴功能的实现; 代码态模式下, 本研究主要借助百度提供的文心大模型接口, 辅以代码完成 AI 中文语伴功能的开发, 低码态开发与代码态开发的流程如图 1 所示。

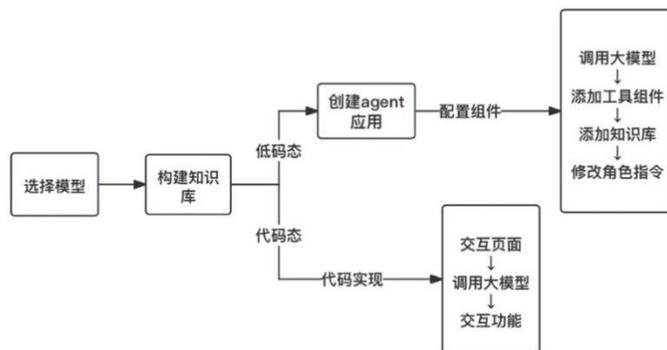


图 1 低码态与代码态的开发流程

第三, 在使用两种开发模式分别搭建完成 AI 中文语伴后, 将两种模式的开发流程及其生成内容进行对比, 并邀请一线国际中文教师和中文学习者分别评价两种 AI 中文语伴的生成效果, 最后给出对于两种开发模式的择用建议。

3. 低码态与高码态模式下的 AI 中文语伴构建流程

3.1 开发准备

3.1.1 模型选择

目前, 国外已经开发了多种专门用于教育领域的人工智能大模型, 如 Khanmigo^①、Merlyn Mind^②和 ChatGPT Edu^③等, 能够解决一些教育场景下的任务需求。国内, 王鑫雷等人 (2020) 对比了目前不同预训练语言模型, 发现百度的 ERNIE (Wang & Fen, 2022) 相较于其他模型在中文问答任务中效果最佳。在中文处理方面, 国内大模型显然比国外的大模型更具有优势, 文心大模型 4.0 能够更好地理解中文语境和

^① Khanmigo 信息参见: <https://www.khanmigo.ai/>

^② Merlyn Mind 信息参见: <https://www.merlyn.org/>

^③ ChatGPT Edu 信息参见: <https://openai.com/index/introducing-chatgpt-edu/>

语义, 在中文处理方面具有较高的准确性, 故本研究选择基于文心大模型搭建 AI 中文语伴。

3.1.2 领域知识库构建

文心大模型为通用型大模型, 能适应多种场景, 但是在特定领域生成的内容存在准确性不高, 专业性不足的问题, 因此需要引入适配微调帮助模型更好地适应特定需求(代必芳&陈贵平, 2024)。由于面向教育类应用, 为增强产品的教育性与科学性, 本研究收集了《新汉语水平考试大纲》、《HSK 标准教程》、HSK 一至六级考试真题与答案解析等资料, 并采用爬虫技术获取了网络上公开发布的国际中文教育文本数据(如国际中文教材、习题、教案等), 实现了领域知识库的构建, 作为领域知识上传至大模型中, 实现国际中文教育领域知识增强, 使模型生成的答案更加科学、精准。

3.2 系统功能设计

本研究拟搭建的 AI 中文语伴预期实现的功能包括三类: 社交功能、教学功能与辅助功能, 具体阐释如下:

3.2.1 社交功能

人机对话功能: 这是最基本的功能。AI 中文语伴应当具备自然语言处理能力, 可以与学习者进行流畅的对话。这不仅限于简单的问答, 还包括进行深入的讨论和交流。此外, AI 中文语伴还要根据历史记录识别用户的语言特点和表达习惯, 从而更好地理解用户的意图, 提供更准确的回复和建议。

情感交互功能: AI 中文语伴与搜索引擎的不同就在于 AI 中文语伴能够提供更加人性化的回答。AI 中文语伴应当具有情感表达能力, 可以从学习者的输入中感知学习者的中文学习过程中的情绪变化, 并据此给予不同的回应, 与学习者建立情感连接, 增强学习者的学习体验。

3.2.2 教学功能

知识问答功能: AI 中文语伴应具备回答中文学习者提出的各种问题的能力, 准确、及时地进行中文知识问答服务, 帮助中文学习者解决各种疑问和困惑。并根据反馈, 提供更加全面、精准的答案。

个性化学习支持功能: AI 中文语伴应该可以根据学习者的学习需求和中文水平, 为学习者推荐适合的学习资源和学习方法, 帮助学习者建立个性化的学习计划, 并提供针对性的学习建议和指导, 从而帮助学习者提高学习效果。

3.2.3 辅助功能

多语种支持功能: AI 中文语伴应具备多语种翻译能力, 实现中文与其他语种之间的即时翻译, 允许学习者需要的情况下借助其它语言理解中文。

3.3 基于低码态模式的 AI 中文语伴开发

基于本研究的研究目标, 本研究使用百度智能云千帆 AppBuilder 智能体构建平台^④提供的服务搭建 AI 中文语伴。该平台为开发者提供多种开源的应用代码框架, 本文选择 Agent Builder 为基础搭建 AI 中文语伴。

低码态的开发包括“基础模式(零代码)”和“工作流模式”, 基础模式创建只需要开发者进行提示词的编写, 工作流模式创建则还需要开发者进行流程设计。低码态下两种创建模式的开发页面如图 2、3 所示。

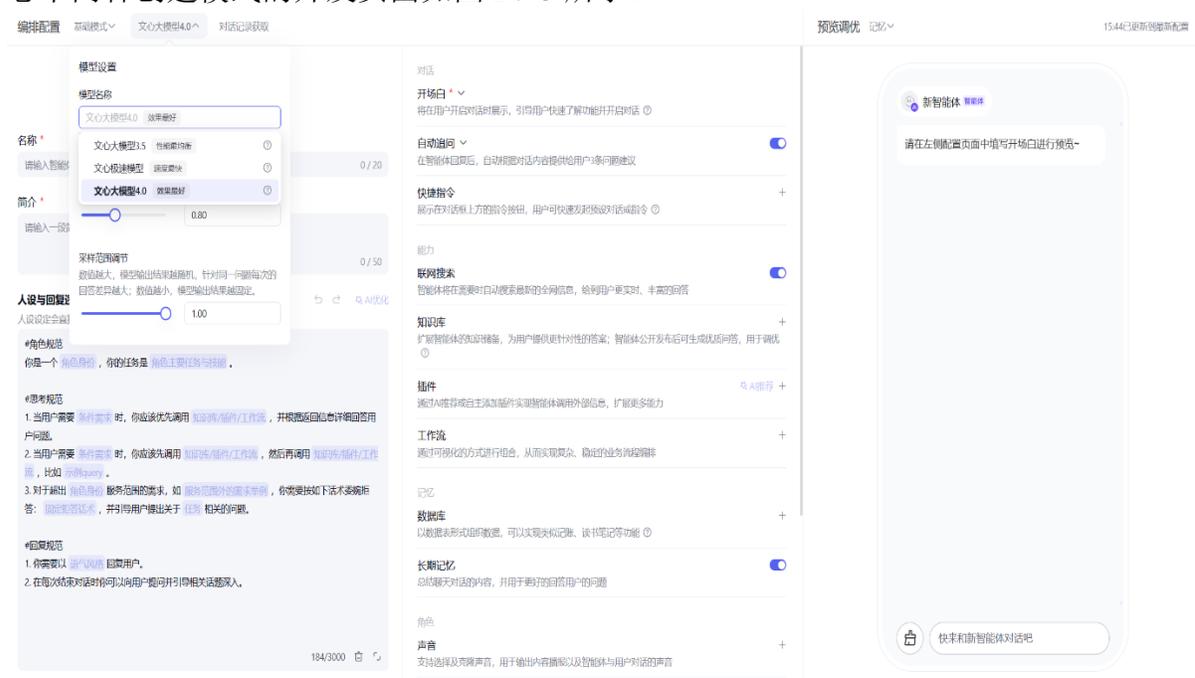


图 2 “零代码模式”开发界面

^④ 平台信息参见: <https://appbuilder.cloud.baidu.com>

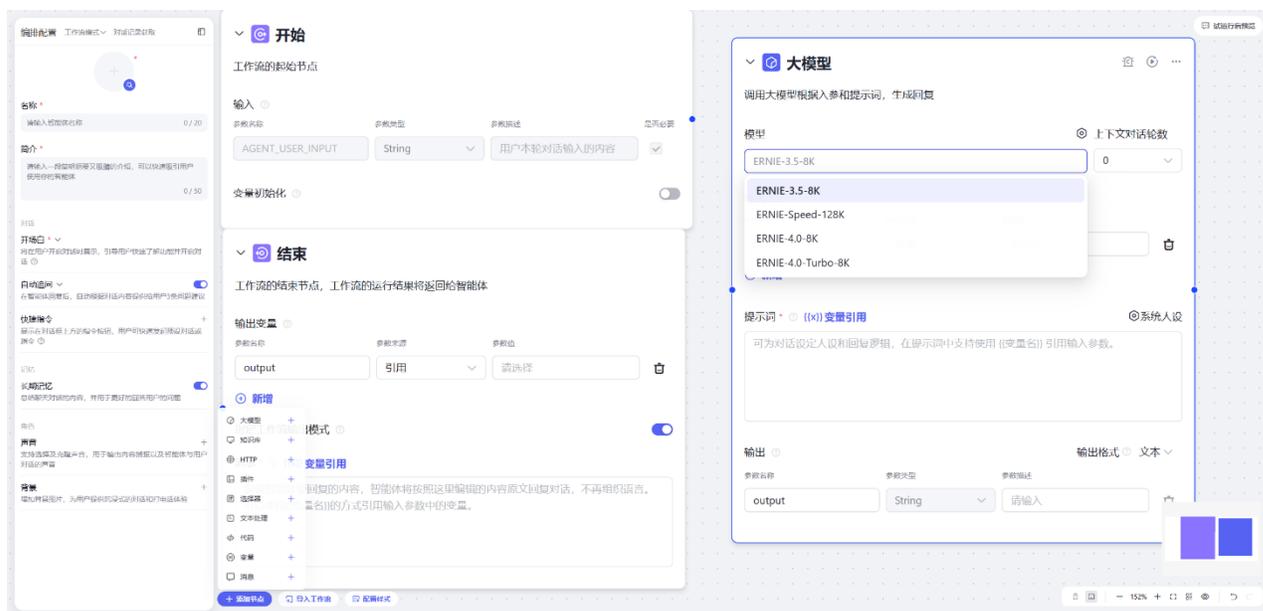


图3 “ workflow 模式”开发界面

本研究中搭建的 AI 中文语伴基于 ERNIE-4.0-8k（文心一言大模型 4.0），模型默认“多样性调节”参数为 0.8、“采样范围调节”为 1.0。为了控制生成内容适合学习者，本实验对两个参数进行了测试，发现当“采样范围调节（随机）”为 1.0 时，每次生成的回答差异性略大，故将该值降低为 0.8，测试发现将采样范围下调为 0.8 后的 AI 中文语伴多次生成的效果更稳定，更符合学习者的学习需求，因此将两个参数均设置为 0.8 进行实验。

本研究选用低码态模式下的 workflow 模式进行 AI 中文语伴开发，workflow 允许用户精细控制每一步的逻辑和输出，这一能力提高了 AI 应用的稳定性和可复现性。通过 workflow 的搭建，可以让智能体更高质量的处理复杂任务。首先对 AI 中文语伴的基本信息、开场白与推荐问题进行设置，说明该 AI 中文语伴的使用场景；接着，针对研究目标，对角色指令做出修改，明确其作为中文 AI 中文语伴陪伴中文学习者进行中文学习的定位，并丰富其功能，如纠正语法错误、提供学习资源与建议、设计互动游戏和挑战等；然后，在“知识库”处上传包含国际中文教育领域知识的文档，使 AI 中文语伴的回答更加专业。

3.4 基于代码态模式的 AI 中文语伴开发

百度提供了多种基线模型供开发者选择，如 ERNIE-3.5-8k、ERNIE-4.0-8k、ERNIE-Speed-128K 等文心一言大模型（不同模型的选择可见图 2、图 3）。为了与低码态下开发的 AI 中文语伴对比更加严谨、可信，通过代码态模式搭建 AI 中文语伴时仍使用低码态搭建 AI 中文语伴时使用的大模型：ERNIE-4.0-8k。

本研究首先通过百度提供的接口将文心一言大模型接入搭建的 AI 中文语伴, 并将收集到的数据全部上传知识库。交互方式上, 本研究借助 Python 内置的 Tkinter^⑤ 库搭建了一个能与中文学习者进行交互的聊天窗口。

在代码态下, 搭建 AI 中文语伴时必须设置的参数值在编写代码时为默认值, 中文学习者能够进行更多的个性化选择, 根据学习需求对参数进行修改。在代码态开发的 AI 中文语伴中, “随机”“多样”分别代表低码态下的“采样范围调节”参数和“多样性调节”参数, 二者均同上设置为 0.8。“惩罚”通过对已生成的 token 增加惩罚, 减少重复生成的现象, 值越大表示惩罚越大, 即更大程度地避免生成重复的 token, 设置为 1.0; “流式输出”提供了更快的响应时间和更低的延迟, 从而提供更好的使用体验; “人设”并非必须设置的参数, 学习者可以按需灵活设置。对于不了解这些参数作用的学习者, 也可以使用默认参数。聊天窗口中参数设置的示例见图 4。



图 4 代码态的参数设置

3.5 开发过程对比

通过对比使用低码态与代码态两种模式的开发全过程, 本研究发现两种开发模式差异如下:

第一, 在对开发者的数字素养要求方面, 由于低代码平台提供可视化的界面和拖放式设计工具, 低码态开发对开发者编程知识要求较低。只要开发者对于应用功

^⑤ 具体文档参考: <https://docs.python.org/zh-cn/3/library/tk.html>

能有具体构想, 借助平台提供的工具, 即使是对代码一窍不通的人也能迅速做出属于自己的 AI 中文语伴。而代码态开发需要开发者具备深入的编程知识和技能。

第二, 在开发效率方面, 由于平台提供了可视化的界面和预制组件, 减少了编写代码的时间, 低码态开发通常能够更快地构建应用程序。

第三, 在页面美观度方面, 低码态使用网站提供的页面形式, 已经满足 AI 中文语伴的外观需求。但代码态的页面美化工作需要开发者自行设定且操作复杂。

第四, 在功能的灵活性和定制化方面, 由于开发者可以根据具体需求编写自定义的代码, 代码态开发通常能够提供更高的灵活性。而低码态开发受到平台提供的预制组件和界面设计工具的限制, 虽不同平台能做出的 AI 中文语伴也有所不同, 但定制化程度相对较低。

第五, 在后续的可维护性方面, 由于低码态开发时使用的是平台封装好的组件, 操作简单, 可维护性强。而代码态开发因为开发者需要自己去定义代码结构和逻辑, 维护起来相对复杂。

第六, 在开发模式适用场景方面, 低码态开发适用于快速迭代或简单的应用程序, 而代码态开发适用于复杂或需要高度定制化的应用程序。

综上, 两种开发模式的对比见表 1, 这两种开发模式各有优缺点, 选择哪种模式搭建 AI 中文语伴取决于具体需求、时间、资源和开发人员的技能水平。

表 1 低码态开发与代码态开发模式的过程对比

项目	低码态模式	代码态模式
数字素养要求	低	较高
开发效率	高	低
美观度	较高	定制化
灵活性和定制化	较差	高
可维护性	高	较高
适用场景	简单、同质化	复杂、定制化

本研究经过上述开发实践, 得出如下结论: 低码态开发的优势在于能够快速搭建出基本功能并获取反馈, 因此, 低码态开发更适用于初期构建 AI 中文语伴。然而, 随着 AI 中文语伴的进一步发展, 它需要更高度的定制化和功能扩展。此时, 代码态开发模式更为合适。代码态开发拥有更好的灵活性和定制性, 能满足特定的需求, 并针对性能瓶颈进行调优。此外, 由于 AI 中文语伴是一个长期发展的项目, 因此可维护性至关重要, 即使对于代码经验相对较弱的国际中文教师而言, 低码态开发模式同样能够提供较好的可维护性, 以适应未来的需求变化和技术更新。因此, 在开发 AI 中文语伴时, 国际中文教师可以考虑从低码态开发开始, 快速验证概念和构建基本功能, 然后逐步转向代码态开发, 以实现更高度的定制化和性能优化。

4. 两种开发模式下 AI 中文语伴的生成效果评测

本研究从两个维度对两种开发模式下构建的 AI 中文语伴进行评测, 首先, 从社交功能、教学功能、辅助功能三个方面对 AI 中文语伴进行基本功能的评测; 其次, 对低码态与代码态两种开发模式下 AI 中文语伴的生成效果进行评测。

4.1 基本功能评测

在社交功能方面, 两种模式下的 AI 中文语伴均能够实现与学习者进行流畅的问答, 并且在回答风格上具有“人情味”, 温柔地指出了学习者的语法错误, 并根据学习者的输出内容进行提问, 鼓励学习者进行更多输出, 如图 5 所示。

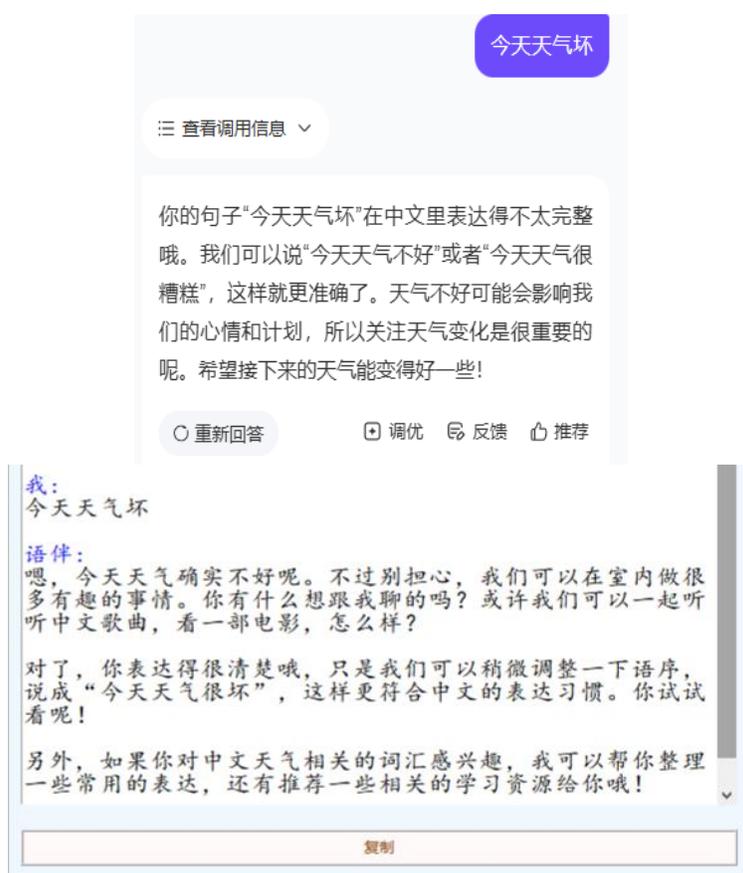


图 5 社交功能展示 (上为低码态, 下为代码态)

在教学功能方面, 两种模式下的 AI 中文语伴均能够妥善回答学习者提出的问题, 并根据学习者的学习目标与时间预算为学习者制定学习计划供学习者参考, 如图 6 所示。

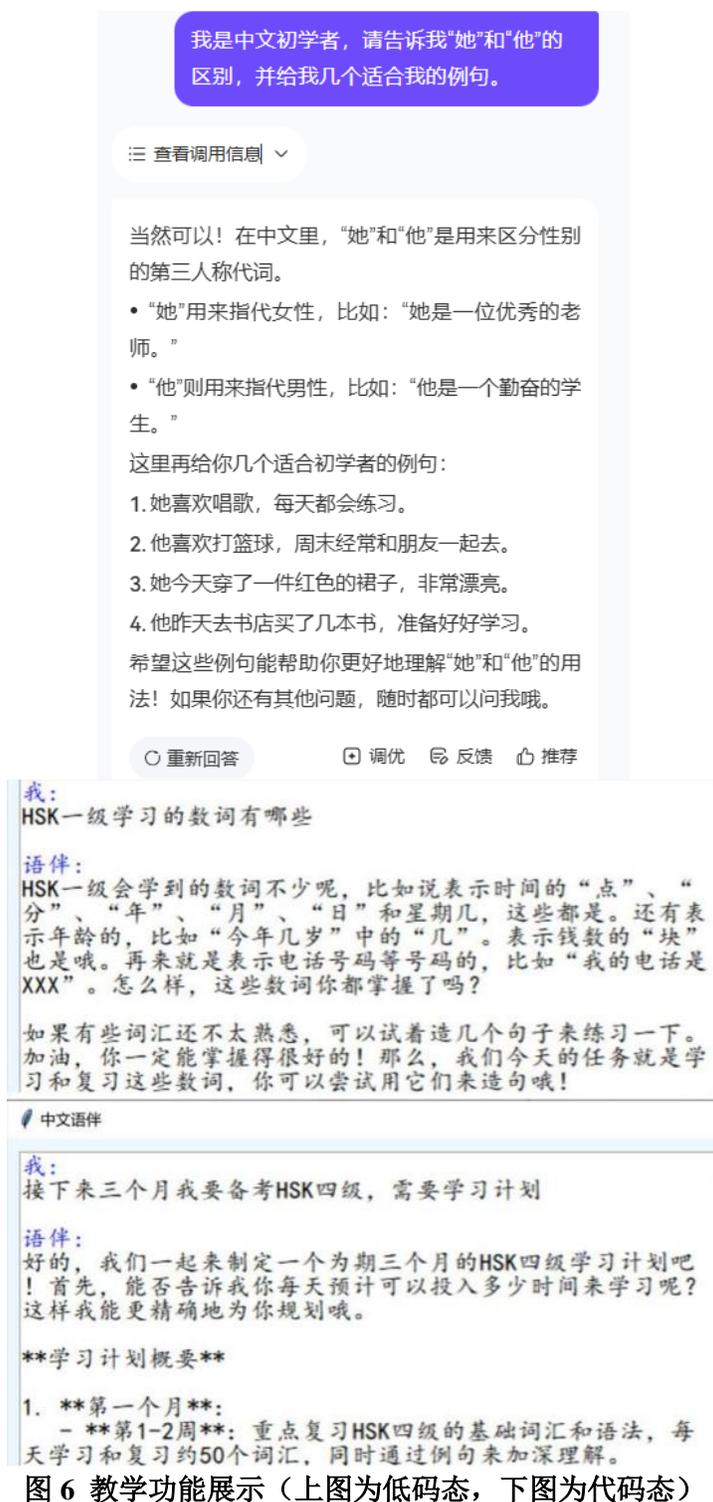


图 6 教学功能展示 (上图为低码态, 下图为代码态)

在辅助功能方面, AI 中文语伴能够根据学习者需求生成文本, 也可以提供中文与学习者母语之间的翻译, 帮助学习者从母语角度理解中文。如图 7 所示, 学习者可以利用 AI 中文语伴辅助学习 HSK 考试的作文部分。当学习者无法理解 AI 中文语伴的回答时, 可以要求 AI 中文语伴将回答翻译为学习者更加熟悉的语言。

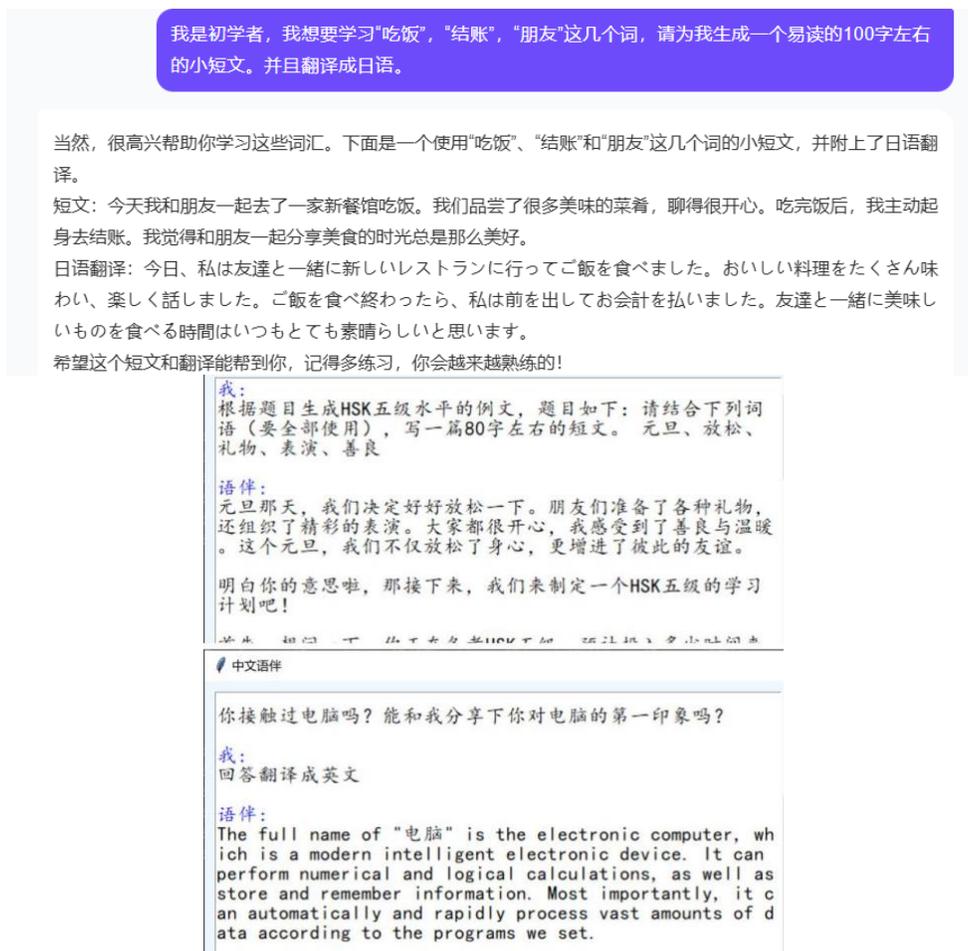


图7 辅助功能展示(上为低码态, 下为代码态)

综上所述, 两种开发模式下开发的 AI 中文语伴均具备 AI 中文语伴的基本功能, 且能够胜任作为 AI 中文语伴的基本任务, 进而可以对开发过程、生成效果进行对比。

4.2 生成内容分析

为进一步检验不同开发模式下研制的 AI 中文语伴在与中文学习者互动中的有效性差异, 本研究调研了三类主要问题场景, 并利用构建好的 AI 中文语伴多次生成答案进行评测分析。本研究面对三类问题场景, 各生成了 10 份不同问题的答案文本, 如分别询问不同等级 HSK 考试的备考建议等, 共计 30 份。由于篇幅原因, 仅展示每类问题的一份答案文本, 具体问题设置和 Prompt 说明如下 ([input]均为用户输入的问题, 生成效果分别见表 2、3、4):

第一类为“建议类问题”, 旨在模拟学习者寻求备考建议的情境, 例如问题: “我现在正在备考 HSK4 级, 你能给我一些备考建议吗?” 在该情境中, Prompt 设置围绕 HSK 备考策略展开, 以引导 AI 生成与考试相关的建议和复习计划, 如“[input],

请给我一些学习建议。请根据我的学习需求, 提供合适的学习计划、复习重点或方法, 帮助我更有效地进步。”

第二类为“对话类问题”, 模拟学习者与 AI 进行人机对话的交流需求, 例如问题: “我们能一起聊聊‘健身’这个话题吗?” 在此类互动中, Prompt 会设置为开放式对话模式, 引导 AI 生成较为丰富的交互内容, 适合具备一定汉语交流能力的学习者, 如“我希望[input]。请用简单易懂的中文与我交流, 使用温和的语气, 并提供有关该主题的相关知识或见解, 帮助我提升中文表达和理解能力。”

第三类为“词义辨析类问题”, 帮助学习者理解常见的词义区分, 例如问题示例: “我不理解‘遇见’和‘遇到’有什么区别, 你能帮我解答吗?” 此类 Prompt 专注于汉语词汇语义的准确解释与辨析, 以满足学习者对具体语言知识的深度需求, 如“你是一名专业的中文教师, 在学习中文时, [input]。请帮助我解释它们的具体含义和使用场景, 并给出相关的用法说明, 以便我更好地掌握这些词汇。”

通过内容对比本研究发现, 从生成质量来看, 代码态开发模式研制的 AI 中文语伴在人机交互方面的效果更佳。具体来看上述三类问题的答案质量:

表 2 两种开发模式下第一类问题的生成效果示例

低码态开发模式	代码态开发模式
<p>当然可以啦! HSK4 级是一个重要的里程碑, 需要你在词汇、语法和听力理解等方面都有一定的积累。以下是一些备考建议, 希望能帮助你更好地准备考试:</p> <p>首先, 词汇积累是非常重要的。你可以每天设定一个词汇学习目标, 比如学习 20 个新词汇, 并复习之前学过的词汇。可以使用词汇卡片、手机 APP 等工具来帮助记忆。</p> <p>其次, 语法学习也是必不可少的。HSK4 级会涉及一些较复杂的语法结构, 你需要花时间学习和理解这些语法规则, 并通过练习来巩固。</p> <p>另外, 听力练习也是非常重要的一环。你可以多听一些与 HSK4 级水平相符的听力材料, 比如 HSK 真题的听力部分, 或者一些中文广播、电视节目等。</p> <p>最后, 模拟考试也是很有帮助的。你可以找一些 HSK4 级的模拟题来做, 模拟考试的环境和时间, 这样可以帮助你更好地适应考试节奏。</p> <p>除了这些, 还有一些辅助学习工具和资源可以帮到你。比如, 你可以找一些中文歌曲、电影和书籍来培养语感, 提高听力水平。同时, 也可以参加一些在线课程或者学习社</p>	<p>当然可以! 为了更有效地备考, 你能告诉我你预计投入多少时间来复习吗? 这样我可以帮你制定一个更贴合你实际情况的学习计划。同时, 你想先从哪方面着手呢, 比如词汇、语法还是模拟试题?</p> <p>词汇方面, 你可以通过背单词卡片或者使用一些在线词汇学习工具来积累词汇量。</p> <p>语法方面, 可以重点复习 HSK4 级常考的语法点, 并结合例句进行练习, 以加深理解。另外, 做模拟试题也是很有帮助的, 它能让你熟悉考试形式和提升答题技巧。</p> <p>当然, 除了这些备考方法, 我还可以为你设计一些互动游戏和挑战, 让你在轻松愉快的氛围中提升汉语水平。比如, 我们可以玩“你画我猜”的游戏来练习词汇, 或者进行“语法接龙”挑战来巩固语法知识。</p> <p>此外, 我也非常推荐你利用一些优质的学习资源来提高学习效果。你可以看一些中文电影或者电视剧来提升听力和口语能力, 同时也可以通过听中文歌曲来学习新词汇和表达方式。如果需要的话, 我还可以帮你找到一些适合你的在线课程, 让你能够更系统地学习。</p> <p>这些建议对你有帮助吗? 或者你还有其他</p>

<p>群, 与其他学习者一起交流学习心得和经验。</p> <p>那么, 你目前每天大概有多少时间可以用来备考呢? 这样我可以为你制定一个更具体的学习计划。</p>	<p>关于备考 HSK4 级的问题想要了解吗?</p> <p><u>同时, 关于你当天的学习任务, 我们可以这样安排: 首先, 掌握 30 个左右的新词汇, 并通过例句来加深理解; 其次, 重点学习和练习 5 个 HSK4 级常考的语法点; 最后, 做一套模拟试题来检验学习成果。你觉得这样的安排怎么样呢? 或者你有其他想法也可以告诉我哦!</u></p>
---	---

首先, 对于建议类问题, 代码态开发的 AI 中文语伴与低码态开发的 AI 中文语伴均从词汇识记、语法点积累、备考方法等方面给出了建议, 总体上看建议较为宽泛, 但同时, 代码态开发的 AI 中文语伴还为学习者专门制定了每日的备考和学习计划, 详见上表中问题 1 的回答 (下划线部分), 这是低码态的 AI 中文语伴未提供的内容。

表 3 两种开发模式下第二类问题的生成效果示例

低码态开发模式	代码态开发模式
<p>当然可以呀! 健身是一个很好的话题, 它对我们的身体健康和心理健康都有很大的益处。</p> <p>你平时喜欢做哪些健身运动呢? 是喜欢去健身房锻炼, 还是更喜欢在家做一些简单的运动呢?</p>	<p>当然可以呀! 健身真的是个很棒的话题呢。你平时的健身习惯是怎样的呢? 是喜欢去健身房, 还是更偏爱户外运动呀?</p> <p>其实, 我对健身也挺感兴趣的, 虽然平时工作忙, 但我一直都尽量抽时间去做些运动。我觉得健身不仅能让我们的身体更健康, 还能提升我们的精神状态, 让我们更有活力地面对生活中的各种挑战, 你觉得呢? 你有没有发现健身给你带来哪些变化呀?</p> <p><u>(在分享的过程中, 如果出现语法错误, 我会自然地指出并给出正确的表达, 例如: “哦, 你是说‘我觉得健身让我变得更自信了’吗? 这句话用完整的句子表达应该是‘我觉得健身让我变得更有自信了。’很棒的观点!”)</u></p> <p>我们也可以一起制定一个健身主题的中文学习计划。比如, 每周学习一定数量的健身相关词汇和句型, 然后通过对话练习来巩固和运用所学内容。这样既能提升你的中文水平, 又能让你更深入地了解健身这个话题, 你觉得怎么样呢?</p>

其次, 对于会话类问题, 代码态开发的 AI 中文语伴回复的内容长度更长, 既能回复自己对“健身”的感受和喜爱情况, 为学习者提供健身常用的词汇和表达方式; 同时, 也能鼓励学习者勇敢开口, 并为学习者输出的文本提供语法纠错服务, 如上表中问题 2 的回答 (下划线部分), 这些都为语言学习提供了良好的“脚手架”。

表 4 两种开发模式下第三类问题的生成效果示例

低码态开发模式	代码态开发模式
<p>当然可以啦~“遇见”和“遇到”在中文里确实有一些微妙的区别。</p> <p>“遇见”通常用于描述人与人之间的相遇，带有一种不期而遇、偶然相遇的意味，比如“我在街上遇见了一个老朋友”。它更多地强调了一种情感色彩和意外性。</p> <p>而“遇到”则更广泛地用于描述各种情况下的相遇，不仅限于人与人之间，还可以用于描述遇到事情、问题等。比如“我今天在路上遇到了一起交通事故”。它更多地强调了一种客观事实的发生。</p> <p>所以，简单来说，“遇见”更侧重于人与人之间的偶然相遇，而“遇到”则更广泛，可以用于各种情况。</p> <p>希望这个解释能帮助你更好地理解这两个词的区别哦~如果你还有其他问题或者想进一步讨论，随时都可以告诉我哦!</p>	<p>当然可以呀!“遇见”和“遇到”在中文里确实有一些微妙的差异呢。一般来说，“遇见”更强调的是一种偶然或意料之外的碰面，比如你在街上突然碰到了一个久未见面的朋友，这时候用“遇见”就更合适。而“遇到”则更常用于描述碰到问题或困难等情境，比如“我在工作中遇到了一个难题”。简单来说，“遇见”更偏向于人或事物的碰面，“遇到”则更广泛，可以包括各种情况或问题的出现。希望这样解释能帮到你哦!</p> <p>那么，为了更好地掌握这两个词，你可以尝试造几个句子来实践一下。比如：“昨天我在公园里遇见了一个老朋友。”或者“最近我在学习上遇到了一些困难。”这样运用在实际语境中，相信你能更快地掌握它们的用法。</p> <p>另外，中文里还有很多类似的近义词，如果你感兴趣的话，我可以帮你整理一些常见的近义词辨析资料，这样可以帮助你更深入地了解中文的细腻之处。你觉得怎么样呢?</p>

最后，对于词义辨析类问题，使用代码态模式开发的 AI 中文语伴辨析的准确性更高，在对“遇见”和“遇到”这组易混淆词进行辨析时，可以准确指出“遇见”常与人或事物（具有“能够看见”性质义素）的词搭配，而遇到的语义范围更大，能够与抽象名词（如困难、挫折等）搭配，为二者提供了直接、有效地辨析。与此同时，代码态开发的 AI 中文语伴还能够为学习者提供两个词语的例句，这符合了词汇教学“词不离句”的教学思路，也可以帮助学习者了解近义词进入句子（语境）后的使用方法，使静态的词汇知识能够迁移到动态的语境中，让学习者真正掌握词汇的语用知识。

综上，从本研究用于测试基于不同开发模式效果的三类问题来看，可以发现，两种开发模式下的 AI 中文语伴都能够胜任其语伴角色，但代码态下的 AI 中文语伴可以为中文学习者提供更优质、丰富的回答内容。

4.3 生成效果评价

为了探讨国际中文教师和中文学习者使用不同开发模式下的 AI 中文语伴的实际效果，本研究在北京语言大学邀请了 5 位一线国际中文教师以及 5 位非中文专业的外国留学生对两种 AI 中文语伴的回答效果进行评价，并以访问链接和本地运行程序（压缩包）的形式，分别将低码态和代码态开发的语伴交给受调查者试用体验，并将上述 30 份“提问—回答”对文本作为调查内容，让受调查者在阅读完答案文本

后, 对语伴的生成质量进行多维度评价。需要说明的是, 本研究并未告知被试不同回答内容是属于何种开发模式下的 AI 中文语伴生成的, 进而保证了评价结果的客观性与真实性。本研究参考 Chiang (2023) 提出的文本生成质量评估标准对两种不同模式下 AI 中文语伴生成的答案从回答相关程度、回答多元程度、回答准确程度、回答详细程度、回答帮助程度五个维度进行评价, 表中每个评价维度下, 评分前者为低码态得分, 后者为代码态得分, 评分为 5 分制, 1-5 分分别代表程度: 非常低、比较低、一般、比较高、非常高, 具体评分如表 5 所示。

表 5 两种模式下生成的回答评分

	回答相关程度		回答多元程度		回答准确程度		回答详细程度		回答帮助程度	
	低码态	代码态	低码态	代码态	低码态	代码态	低码态	代码态	低码态	代码态
教师 1	4.5	4.5	4	4	4	4	4	4.5	4	4.5
教师 2	3	4.5	3	4.5	4	4	3.5	4	3	4
教师 3	3.5	4	4	4	3	3.5	3	4	3	3.5
教师 4	4	4.5	3.5	4.5	3.5	4	3	4	3.5	3.5
教师 5	3.5	4.5	3	4.5	3.5	4	3	4.5	3.5	4
学生 1	3.5	4	3	3.5	4	4	3	4	3.5	4
学生 2	4	4.5	3	4	4	4	3	4	3.5	4.5
学生 3	4	4	3	4.5	3.5	4	3	4.5	3	3
学生 4	4	4.5	3	4.5	3.5	4	3.5	4.5	4.5	4.5
学生 5	4	4	3	4	4	4	3	4	4	4
均分	3.8	4.3	3.25	4.2	3.7	3.95	3.2	4.2	3.55	3.95

由上表可知, 通过对比可以发现, 代码态下的回答分数相对更高, 这也说明两种开发模式对生成内容质量会产生一定影响, 主要表现在回答内容的详细程度和有用性方面。进一步看, 在受调查者中, 5 位中文教师均认为通过低码态构建的 AI 中文语伴已符合他们的教学需求, 但代码态的回答在五个维度相对更好, 但同时也有教师提出, 代码态 AI 中文语伴生成的回答过多, 可能会对学习者造成认知负担, 希望可以进一步对代码态 AI 进行调整; 5 位外国留学生也认为低码态下的 AI 中文语伴已满足学习需求, 但有 4 位留学生对构建代码态的 AI 中文语伴充满兴趣, 想要进一步定制化开发专属的 AI 中文语伴。

4.4 开发模式择用建议

经过对两种开发模式不同角度的对比, 本研究发现, 两种开发模式在过程和效果上的优势和不足不尽相同, 总体而言, 低码态开发模式具有低耗时的特质, 低码态开发在软件开发和维护过程中均提高了生产力, 且通过预编程的应用组件和数据模型工具, 简化了开发流程, 对于想要实现基础功能且“即插即用”效果的开发基础薄弱的中文教师是个更优的选择。而相应的, 这种开发模式研制的 AI 中文语伴的效果并不优于代码态开发的产品, 对于想要更高定制化的效果则需引入代码态开发模式。

因此, 未来中文教师或学习者在使用两种开发模式时, 可以根据自身实际需求, 灵活选择最适宜的开发模式, 且由于大部分中文教师未曾受过专业的代码训练, 开发能力较弱, 可以优先考虑使用低码态的开发模式, 这种模式同样可以研发出基本胜任语伴工作的智能体。如果教师希望语伴能够更加个性化、智慧化地发挥其教育价值, 需要不断提升自身的数字素养, 特别是数字教学设计能力、数字技术应用能力(方紫帆 & 徐娟, 2024), 或者选择与擅长开发的技术人员合作, 形成教学资源建设共同体, 运用代码态的开发模式协作构建更针对教学需求的 AI 中文语伴。

5. 结论

目前, 在国际中文教育领域, 中文语伴资源存在明显困境。本文从“低码态”和“代码态”两个角度出发, 选择百度文心一言大模型作为 AI 中文语伴的基线模型, 对 AI 中文语伴从设计到具体实现进行了全面的介绍, 并在搭建完成后对两种开发模式下的开发过程和效果进行对比, 得出两种开发模式的利弊所在, 未来开发实践中应当结合二者优点完成 AI 中文语伴的开发。

尽管本研究在 AI 中文语伴方面取得了一定进展, 但在未来仍有许多改进空间, 例如: 第一, 缺少对比分析其他大模型及开发平台。本研究暂未对比其他先进的大语言模型, 如 Llama、GPT-4、GLM 等, 也未涵盖不同开发平台的比较分析。若纳入更广泛的模型和平台比较, 可以进一步为结论进行验证, 提升研究成果的普适性; 第二, 实验覆盖的问题种类不够全面。本研究的实验应用主要集中于三类问题, 这可能限制了对模型在不同问题背景下的表现评估。尽管所选问题具有重要意义, 但未能涵盖更广泛的问题种类可能导致研究结果的适用性受到限制; 第三, 实验所涉及的样本规模相对较小。本研究仅包含 5 位一线教师和 5 位二语学习者。这可能影响结果的统计显著性和可靠性; 第四, 暂时缺乏长期实际效果评估。本研究未能对所开发的智能体进行长期效果的评估, 虽然初步结果显示了模型的有效性, 但长远来看具体应用效果如何不能确定, 这一局限性可能影响对模型实际应用价值的全面理解。

随着人工智能技术的不断进步, AI 中文语伴将成为中文学习者的重要伙伴和支持者, 为中文学习者带来更广阔的学习空间和更丰富的学习资源, 推动国际中文教育的规模化普及和个性化培养并举。

致谢: 本文受世界汉语教学学会全球中文教育主题学术活动计划“融合知识图谱与大语言模型的国际中文教育智能问答系统研究”(SH24Y24); 北京语言大学研究生创新基金(中央高校基本科研业务费专项资金)“国际中文教育学科知识图谱问答系统的研究与构建”(24YCX019)的资助。徐娟为本文通讯作者。北京语言大学信息科学学院马瑞凌、司函等在论文修改过程中提出了许多宝贵建议, 谨此致谢。

参考文献

- Chiang, C. H., & Lee, H. Y. (2023). Can large language models be an alternative to human evaluations? <https://arxiv.org/abs/2305.01937>
- Chen, Q. (2020). A study on the group discussion of language partner cooperation models in teaching Chinese as a foreign language: A case study of Jiangsu University. *Modern Communication*, 2020(12), 40-41. [陈倩. (2020). 对外汉语语伴合作模式小组讨论研究——以江苏大学为例. *现代交际*, 2020(12), 40-41.]
- Cao, G., Dong, Z., & Xu, J. (2023). Construction of a vocabulary knowledge graph and adaptive vocabulary learning platform based on the Chinese Language Proficiency Standards for International Chinese Education. *International Journal of Chinese Language Teaching Research*, 2023(1), 21-30. [曹钢, 董政, & 徐娟. (2023). 基于《国际中文教育中文水平等级标准》的词汇知识图谱与词汇自适应学习平台构建. *国际汉语教学研究*, 2023(1), 21-30.]
- Dai, B. F., & Chen, G. P. (2024). Research on dialogue and improvement strategies of artificial intelligence question-answering systems: A case study of Wenxin Yiyan. *Computer Knowledge and Technology*, 2024(8), 14-16+38. [代必芳, & 陈贵平. (2024). 人工智能问答系统的对话与改进策略研究——以文心一言为例. *电脑知识与技术*, 2024(8), 14-16+38.]
- Fang, Z., & Xu, J. (2024). Construction and development path of blockchain-based digital literacy micro-certification system for international Chinese teachers. *Journal of Yunnan Normal University (Teaching & Studying Chinese as a Foreign Language Edition)*, 22(2), 27-38. [方紫帆, & 徐娟. (2024). 基于区块链的国际中文教师数字素养微认证系统构建与发展路径探析. *云南师范大学学报(对外汉语教学与研究版)*, 22(2), 27-38.]
- Mizumoto, A., & Eguchi, M. (2023). Exploring the potential of using an AI language model for automated essay scoring. *Research Methods in Applied Linguistics*, 2(2), 100050.
- Ma, R., & Xu, J. (2024). Innovations in the development of intelligent teaching resources for Chinese writing in the digital age. *International Journal of Chinese Language Teaching Research*, 2024(1), 13-23. [马瑞凌, & 徐娟. (2024). 数智时代国际中文写作智慧教学资源创新研发. *国际汉语教学研究*, 2024(1), 13-23.]

- Rokis, K., & Kirikova, M. (2022). Challenges of low-code/no-code software development: A literature review. In *International conference on business informatics research* (pp. 3-17). Springer International Publishing.
- Sheng, J., Xu, C., & Zhou, T. (2024). Analysis of Python web scraping technology. *China Information World*, 2024(6), 210-212. [盛景, 徐超, &周涛. (2024). Python 网络爬虫技术分析. *中国信息界*, 2024(6), 210-212.]
- Trigo, A., Varajão, J., & Almeida, M. (2022). Low-code versus code-based software development: Which wins the productivity game? *IT Professional*, 24(5), 61-68.
- Wang, X. L., Li, S. C., Yang, Z. H., Lin, H. F., & Wang, J. (2020). Chinese knowledge graph question answering system based on pre-trained language model. *Journal of Shanxi University (Natural Science Edition)*, 43(4), 955-962. [王鑫雷, 李帅驰, 杨志豪, 林鸿飞, &王健. (2020). 基于预训练语言模型的中文知识图谱问答系统. *山西大学学报(自然科学版)*, 43(4), 955-962.]
- Wang, C., & Feng, F. (2022). ERNIE based intelligent triage system. *Journal of Intelligent & Fuzzy Systems*, 43(4), 5013-5022.
- Wooldridge, M., & Jennings, N. R. (1994). Agent theories, architectures, and languages: A survey. *International Workshop on Agent Theories, Architectures, and Languages* (pp. 1-39). Springer Berlin Heidelberg.
- Yan, Y., Guo, X., Wang, T., Rao, Q., & Wang, K. (2023). A review of question-answering systems. *Computer Systems Applications*, 2023(8), 1-18. [闫悦, 郭晓然, 王铁君, 饶强, &王铠杰. (2023). 问答系统研究综述. *计算机系统应用*, 2023(8), 1-18.]
- Zhang, C. H., Du, L. F., Zhu, X. N., & Zhao, H. (2023). Research on education question-answering systems based on large language models. *Journal of Beijing University of Posts and Telecommunications (Social Sciences Edition)*, 2023(6), 79-88. [张春红, 杜龙飞, 朱新宁, &赵慧. (2023). 基于大语言模型的教育问答系统研究. *北京邮电大学学报(社会科学版)*, 2023(6), 79-88.]
- Zheng, Y. H., Wang, C. F., & Zhang, W. W. (2024). Application and regulation of generative artificial intelligence in education. *China Educational Technology*, 2024(5), 114-119. [郑永红, 王辰飞, &张务伟. (2024). 生成式人工智能教育应用及其规制. *中国电化教育*, 2024(5), 114-119.]
- Zhou, X. Y. (2024). A brief discussion on the role of language partner relationships in improving the language proficiency of international students. *Teaching and Research of National Common Language and Characters*, 2024(8), 25-27. [周欣宜. (2024). 浅谈语伴关系对于提高留学生语言水平的作用. *国家通用语言文字教学与研究*, 2024(8), 25-27.]