心理科学进展 2026, Vol. 34, No. 7, 1284–1298 https://doi.org/10.3724/SP.J.1042.2026.1284
Advances in Psychological Science https://cstr.cn/32111.14.2026.1284

·研究方法(Research Method)·

人工智能在心理传记学中的应用可能、挑战与启示*

舒跃育 李春江 任笑笑 谢 霞 张银霞 宋 欢

(西北师范大学心理学院, 兰州 730070)

摘 要 人工智能领域大语言模型技术的快速发展, 为心理传记学应对以往研究中数据量大、处理复杂、主观性强等挑战带来了机遇。本文通过具体案例系统探讨了大语言模型在心理传记学研究中的应用可能性、挑战与启示。在应用可能性方面, 大语言模型可以提高研究效率。但同时也带来数据偏见、解释偏差及语义理解局限等潜在问题与伦理风险。面对这些挑战, 本文提出了三点可能的解决路径: 第一, 构建本土化大语言模型资料库, 为应对文化语境与数据偏见问题提供解决方案; 第二, 提出心理传记领域的人机协同伦理准则, 为人工智能时代心理传记研究的规范发展提供指引; 第三, 提出构建以研究者为主导的人机协同研究框架的理论构想, 为心理传记学在人工智能时代的方法论路径指明方向。

关键词 人工智能, 心理传记学, 大语言模型, 人机协同, DeepSeek

分类号 B841

1 引言

心理传记学是质性研究中关注典型个体的代表性方法(黄希庭, 2021, pp. 127−129), 以探索人性的复杂性与生命的独特性为核心, 旨在通过分析那些遗落在定量研究视野之外的“另类个体”来发现已有理论的不足, 促进心理学理论的完善并推动学科发展(蒋怀滨, 张斌, 2019, p. 313; 舒跃育, 2018b, 2024)。但传统心理传记学研究依赖对海量文本进行手工处理与反复解读, 耗时耗力, 令许多研究者望而却步。近年来, 随着人工智能技术的快速发展, 大语言模型(Large Language Models, LLMs)“已经在语义‘理解’、知识抽取、‘发现’隐含的逻辑关系以及知识归纳方面, 表现出了强大的处理能力” (舒跃育, 2024), 有望在提升编码效率、支撑理论创新等方面发挥关键作用(颜文靖, 2024)。在这一技术背景下, Mayer 和 Fouché (2025)提出了心理传记 4.0 的概念, 主张心理传记研究必须跟随时代发展潮流, 发扬心理传记学在人工智能时代的学科价值。那么大语言模型如何在不损害心理传记学学科本质的前提下, 促进该领域的发展? 具体而言: (1)在以“悬疑性问题”为核心的心理传记学分析模式中, 大语言模型可以在哪些步骤为研究者提供有效支持? (2)在应用于各步骤时, 存在哪些潜在的风险与挑战? (3)为应对这些风险和挑战, 应构建怎样的伦理准则、研究范式或指导框架? 要回答这些问题, 首先需要考察大语言模型与质性研究的结合带来了哪些机遇和挑战。

2 大语言模型为质性研究带来的机遇和挑战

质性研究方法作为一个伞概念(umbrella concept), 包含扎根理论、话语分析、生命史与心理传记等许多不同的研究方法(何吴明, 郑剑虹, 2019)。这些异质方法都以研究者的深度参与、批判性思维和理论创新为核心(傅安国, 岳童, 2024), 关注人们自身体验中的意义获得, 最终帮助研究者在特定文化历史背景下深入理解焦点群体和典型个案(叶浩生, 2009; 周明洁, 张建新, 2008)。

目前已有很多研究者尝试探索运用大语言模型开展各项质性研究(Gibson & Beattie, 2024;

收稿日期: 2025-10-29

* 教育部哲学社会科学研究重大专项“中华民族优秀文化心理学思想研究” (项目号: 2025JZDZ023)。

通信作者: 舒跃育, E-mail: shuyueyu@nwnu.edu.cn

表 1　大语言模型在不同研究环节的应用情况示例

应用环节	应用案例	应用效果	潜在挑战与风险
资料收集与预处理	数据熟悉与摘要生成(Anakok et al., 2025); 转录与翻译(Barrera et al., 2025)	自动化效率高, 但对文化语境的适应性不足, 缺乏人类深度的语境理解。	语义理解局限; 数据偏见
编码与主题分析辅助	自动编码(Tschisgale et al., 2023); 识别与生成主题, 主题审查(Anakok et al., 2025)	在描述性主题识别、关键词提取和信息初筛方面准确率较高, 能够提升编码效率, 但容易忽略边缘性语境细节, 需人工审核。	数据偏见; 透明度不足; 研究者主体性削弱
理论生成、验证与拓展	识别人类研究者可能忽略的模式(Hamilton et al., 2023); 基于特定理论视角重新审视数据, 激发理论对话(Gur & Maaravi, 2025)	在模式发现和基于已有理论的推演上表现出潜力, 但自主生成原创、系统化理论的能力仍有限。	认识论冲突; 去人性化风险; 责任归属模糊; 降低理论敏感性
写作与结果生成	研究结果综合与撰写(傅安国, 岳童, 2024); 代码书编制(Perkins & Roe, 2024)	在遵循学术规范的结构化、标准化文本生成方面效率较高。但生成的文本往往流于表面综合, 缺乏真正的问题意识、论证张力和个人化的学术声音。	学术诚信与原创性危机; 叙事同质化; 伦理敏感性缺失; 人类主体性丧失

Hamilton et al., 2023; Morgan, 2023; Smirnov, 2025; Zhang et al., 2023) (见表 1)。在资料检索和预处理方面, 大语言模型被用于数据转录、内容翻译等基础性预处理工作, 以减轻研究者的负担(Barrera et al., 2025; Chen, 2023; Chopra & Haaland, 2023)。在编码辅助方面, 通过设计特定提示框架(Zhang et al., 2023)或将经典分析方法流程化(Anakok et al., 2025), 能进一步提升大语言模型在识别描述性主题方面的准确性(Brondani et al., 2024), 并与人类研究者互为补充(Hamilton et al., 2023)。但也有研究者认为大语言模型对深层隐含主题的解读仍存在局限(Morgan, 2023)。在理论生成与验证方面, 大语言模型可以基于海量数据资源和跨领域知识库, 帮助研究者发现尚未考虑到的模式, 增强结果的可靠性(Hitch, 2024)。但也有学者认为, 需警惕其可能推动方法论向实证主义范式回潮的结构性风险(Chatzichristos, 2025)。在写作与结果生成方面, 大语言模型已能协助生成研究初稿、撰写文献综述(傅安国, 岳童, 2024), 并通过与传统方法比较, 辅助完成编码手册的编制等工作(Perkins & Roe, 2024)。但有研究表明, 研究者对“使用 AI 撰写论文”的接受度和伦理认可度在所有功能中最低, 审稿人同样对作者使用AI写论文的做法持最消极的态度(Marshall & Naff, 2024), 这表明在多大程度上使用 AI 面临争议。

总体而言, 大语言模型的融入带来了显著的效率优势, 如节省时间、减轻编码强度(De Paoli, 2024; Mathis et al., 2024), 并能处理人力难以应对的大规模数据集(Tschisgale et al., 2023)。但也带来了价值观偏见(Liang et al., 2021; Navigli et al., 2023)、人类主体性丧失(Marshall & Naff, 2024)等风险和挑战。当未来的科研生态不可避免地向“AI——人类协作模式”迁移时, 我们面临的问题不再是“能不能用 AI”, 而是“如何更高质量地用 AI” (He & Bu, 2026)。在这一趋势下, 心理传记学作为质性研究中致力于深度诠释个体生命独特性的代表性方法(黄希庭, 2021, pp. 127−129), 如何在利用智能工具提升研究效率的同时, 坚守其基于深度人性理解与叙事整合的学科内核? 下文将以“悬疑性问题”心理传记分析模式为例, 具体探讨大语言模型在其中的优势、挑战与启示。

3　大语言模型如何推进心理传记学研究

在以“悬疑性问题”为核心的心理传记分析模式中, 包括选择传主、提出悬疑性问题、资料的选择与分析、解释理论的选择与发展以及研究的有效性考察和论文撰写几部分(舒跃育, 2018a, pp. 43−48)。下文将以谢霞(2022)对李清照的研究——《女性真实自我的寻求：以李清照的心理传记学分析为例》一文为主要参考例文, 通过大语言模型(相关版本等信息见附录 A)复现该文的研究步骤, 以探究其在心理传记学研究中的优势与面临的挑战[1]。

[1] 本文中所有运用大语言模型进行的分析, 不同于以往实验研究中的实证分析模式, 而属于为本研究提供论据的“启发式示范”。

3.1 快速定位备选传主

选择一位传主是心理传记学研究的第一步(舒尔茨, 2005/2011, p. 55)。心理传记学中的传主主要指对历史进程产生重要影响的非凡人物, 包括历史人物和在世人物。传主的选择可从两个维度展开, 共包含四个方面：一是传主数量为单个还是多个; 二是选择依据是基于对人物的熟悉程度, 还是基于对某一理论视角的兴趣。如陈雪仪(2021)对林徽因的研究是基于对人物的熟悉度, 而王泽钰(2022)对武则天和慈禧的研究则是基于其女性主义的理论视角。当研究者基于对人物的熟悉度选择传主时, 并不需要大语言模型的参与, 因为传主已经确定。而基于理论视角选择传主时, 大语言模型便可以帮助研究者初步筛选出与该理论相关的传主。例如, 我们想要寻找与“真实自我”理论相关性较高的传主, 通过将要求输入大语言模型, 其给出了辜鸿铭、维特根斯坦、张爱玲等人物。这种从海量数据中进行筛查的能力正是大语言模型所擅长的。检索增强生成(Retrieval-Augmented Generation, RAG)技术可以“将传统的语言生成模型与外部知识源相结合, 提升大语言模型的检索能力” (秦小林 等, 2025), 从而根据研究者的标准, 从海量信息中初步筛选出潜在的研究对象。研究者可在此基础上进一步核实这些人物的成长经历, 以判断其是否适合作为研究的传主。尽管其给出的结果存在一定误差, 但无疑大大减少了研究者从无到有确定传主的工作量。

然而, 大语言模型这种基于海量数据与模式匹配的推荐能力, 在心理传记学的传主选择上存在固有的认识论局限。首先, 其推荐机制本质上是“回溯性”和“强化主流”的。模型倾向于推荐那些在训练数据中被高频讨论、史料丰富的人物, 如辜鸿铭、张爱玲等。这无形中巩固了既有的历史名人谱系, 而可能将那些同样极具心理研究价值, 但因其边缘地位或史料散佚而未被充分文本化与数字化的“异质个体”排除在视野之外。例如, 一位在地方志中仅有零星记载, 却以独特方式反抗时代性别规范的女性, 很可能无法被大语言模型捕获。其次, 大语言模型的“高效”可能带来一定风险：它将研究者的探索, 从一种充满个人洞察的“主动追寻”, 部分地转变为对算法给出的、符合公共叙事的清单的“被动筛选”。因此, 研究者必须警惕, 避免自身的研究目的被技术的可见性框架悄然塑造。

3.2 高效收集可能资料

在初步确定传主后, 需要收集与该传主相关的资料并进行分析处理。在这一阶段, “大语言模型可以从网络上抓取并储存与传主相关的大量公开资料” (潘晓英 等, 2020)。如以宋高宗为例, 大语言模型的应用可贯穿从资料普查到文本倾向性评估的全过程。首先, 在资料收集层面, 研究者仅需输入基本指令, 模型便能依托其内嵌的广泛知识库与检索增强生成(RAG)技术, 快速生成一份结构化的资料索引清单(如表 2 所示)。这份清单系统涵盖了传主所处的宏观社会背景、核心史料以及当代重要的研究专著, 能够迅速勾勒出研究的资料版图, 极大提升研究者评估资料丰富性与可获得性的效率, 使其得以将精力从漫无目的的泛读转向目标明确的精读与深度验证。

表 2 DeepSeek 给出的宋高宗相关资料

类型	书名	作者/编者	出处
史料	《宋史》	脱脱等(元官修)	中华书局(1977)
	《建炎以来系年要录》	李心传(宋)	中华书局(2013 重印)
	《三朝北盟会编》	徐梦莘(宋)	上海古籍出版社(2019)
	《鄂国金佗稡编续编》	岳珂(宋)	中华书局(1989)
研究专著	《岳飞新传》	王曾瑜	中国书籍出版社(2016)
	《宋高宗传》	王曾瑜	中国书籍出版社(2016)
	《秦桧研究》	韩酉山	中华书局(2015)
	《南宋初期政治史研究》	[日]寺地遵	复旦大学出版社(2016)
	《忧患与悲情：南宋初年的时代氛围与文人心态》	郭学信	中国社会科学出版社(2021)
	《皇权下的通权达变：宋高宗心理传记》	何勇强	浙江大学出版社(2022)
	《宋光宗宋宁宗》	虞云国	吉林文史出版社(1997)

更进一步, 在资料初步整理之后, 大语言模型可借助模式识别(Pattern Recognition)与立场检测(Stance Detection)等算法, 对具体文献的内容倾向性进行初步判断(尚钰 等, 2024)。例如, 运用 DeepSeek 分别对王曾瑜(2016)的《宋高宗传》和何忠礼(2021)的《宋高宗新论》进行模式识别和立场检测, 其对两书的分析结果与笔者的阅读感受和学界评论(曹家齐, 2022)基本一致(详细内容见附录 C)。这一功能为研究者在汗牛充栋的文献中筛选相对客观、平衡的论述提供了有价值的初步参考。

然而, 必须清醒地认识到, 对这种高效工具的依赖伴随双重风险, 要求研究者始终保持批判性审视。在“广度”层面, 尽管模型生成的资料清单极具启发性, 但其本质上是对既有数字化、文本化信息的聚合与推断, 可能存在“虚构书目”或遗漏关键稀有文献的风险(相关问题探讨参见附录 B)。在“深度”层面, 模型进行立场检测所依赖的判断标准与训练数据, 不可避免地内嵌了现有的学术共识、主流叙事乃至潜在的数据偏见。这意味着, 模型所谓的“客观”分析本身可能是经特定视角过滤后的产物, 其结论可能在无意中强化主流论述, 而对非主流或批判性视角缺乏敏感性。因此, 大语言模型在这一阶段所提供的, 是一个需要研究者运用专业知识进行严格核验、补充与批判性修正的参考资料库。它将研究者从信息的“搬运工”角色中解放出来, 但对其成为更精明的“鉴定者”与“分析师”提出了更高要求。

3.3　辅助提出悬疑问题

在确定传主以后, 如何提出一个具有张力的“悬疑性问题”, 是心理传记研究的核心与难点。这需要研究者深入把握传主生命中矛盾、反常与震撼之处, 或是贯穿传主大部分生涯的“弥散性、渗透性”的问题(郑剑虹, 2013;舒跃育, 2018a, p. 45; 谢霞, 2022)。大语言模型借助于主题建模(Topic Modeling)、注意力机制(Attention Mechanism)、矛盾检测(Contradiction Detection)等技术, 可以对传主资料进行多维度解析, 发现其中隐含的异常模式(任欢, 王旭光, 2021; 张培晶, 宋蕾, 2012)。

以李清照为例, 运用 DeepSeek 快速读取资料的能力对相关资料进行分析。主题建模技术可以通过对资料的分析进行主题聚类, 识别其人生不同阶段的主题密度变化与中断线索。如再嫁前的词作主题以“惜春”、“伤春”为主, 再嫁后以“分离”、“亡国”为主。注意力机制能够发现读者对“赵明诚临终行为”高度关注, 但原文仅以“殊无分香卖履之意”一笔带过。“分香卖履”典故的隐含信息被高度关注。矛盾检测技术发现其《钓台》一诗表达的寻求殉国的志向与实际随朝廷南逃之间的矛盾。最终由其总结出李清照一生中在常人眼中显得“不合理”之处(见表 3), 从而为研究者提供提出悬疑性问题的结构化线索。

表 3　DeepSeek 分析的李清照一生中“不合理”之处

类别	具体表现	时代常规与对比
才华的“不合理”	公开出版词集：将个人词作整理为《漱玉词》并公开流传, 进入由男性主导的公共文学领域。 撰写《词论》: 系统性地犀利点评北宋名家(如柳永、苏轼、秦观等), 展现学术自信与批判精神。	女性应“无才便是德”, 创作多为私域表达。 文坛由男性垄断, 女性公开批评大家被视为狂妄失礼。
婚姻的“不合理”	以爱情为基础：与赵明诚的婚姻有深厚情感基础。 精神平等伴侣:共同致力于收藏、学术研究与文学唱和, 关系超越“夫为妻纲” (如“赌书泼茶”的佳话)。	士大夫婚姻多遵循“父母之命”, 情感非首要。 夫妻关系为主从模式, 妻子需服从丈夫, 智力平等对话罕见。
行为的“不合理”	晚年再嫁：年近五旬、名满天下时选择再嫁张汝舟。 诉讼离婚：发现骗婚与家暴后, 不惜触犯“妻告夫, 徒二年”的法律状告丈夫, 宁愿坐牢也要离婚(后仅关押九日)。	士大夫遗孀再嫁被视为失节。 女性提出离婚极罕见, 状告丈夫更属大逆不道, 法律对妻子控告丈夫有严厉惩罚。
爱好的“不合理”	酗酒：词中频繁记录饮酒醉酒(如“沉醉不知归路”“浓睡不消残酒”)。 赌博：自称“性喜博”, 精通“打马”游戏并撰文宣扬(《打马图经序》)。	女性饮酒需谨慎含蓄, 公开书写醉态有违淑女形象。 赌博被视为不良嗜好, 女性公开承认并夸耀此好极为罕见。
晚景的“不合理”	守护文物南渡：战乱中独自护送 15 车金石文物辗转千里, 将文化传承置于个人安危之上。 完成遗愿：在困苦晚年整理校勘《金石录》并作后序, 留下学术与文学名篇。	乱世中常人优先保全性命与财产, 无力顾及笨重藏品。 女性独自承担重大文化使命极为少见,《金石录》整理工作本应由男性学者完成。

然而，必须清醒认识到，大语言模型所识别出的“不合理”清单，本质上是对传主行为与时代常规之间“统计偏差”的事实罗列(例如高效地列出李清照酗酒、赌博、讼夫等违背当时社会规范的行为)，这与心理传记学所追求的“悬疑性问题”存在本质区别。一个真正的“悬疑性问题”并非一系列反常行为的罗列，它要求研究者透过表象，追问这些矛盾如何构成了传主生命的内在统一性与动力核心。例如，李清照的诸多“不合理”，是基于对“士人”身份与生活方式的僭越性追求，还是源于其内在自我的冲突与挣扎？

大语言模型的介入，迫使我们必须从本体论层面澄清“悬疑性问题”的构成。其一，是其客观实在性基础，即源于传主生平与时代背景的历史事实、可观测的行为矛盾及引发的集体困惑；其二，是其诠释学意义上的生命力，即研究者基于自身生命体验、理论素养与文化共鸣而对问题产生的“情感指向”与“存在性投入”。正是后者，赋予了悬疑性问题以独特的“问题性”，使之从一个事实罗列升华为一个值得追问的学术命题。大语言模型作为高效的工具，能够高效处理前者，重组“事实网络”；但对后者的“意义赋予”则无能为力。这定义了人机协同的分工本质：大语言模型拓展研究者认知的广度与处理效率，研究者则肩负起对意义的定位、诠释与理论创造。心理传记学最深层的价值，正体现在这种唯有人类主体方能完成的、从事实到意义的创造性跨越之中。

3.4 跨领域提供备选理论

在完成资料处理与关键信息提取后，研究者需要运用心理学理论将其组织成富有启发性、条理清晰的故事(麦克亚当斯，2009/2024，p. 512)，以对传主的悬疑性问题做出合理的解释。这一过程高度依赖于研究者的理论素养与创造性思维。大语言模型凭借其庞大的知识图谱(Knowledge Graph)，能够快速进行跨理论检索与语义匹配(丁邱 等，2023；钱慎一 等，2025)，为研究者提供超越其熟悉范围的多元化的理论备选方案。

例如，DeepSeek 针对李清照“为何再嫁”的悬疑性问题，给出了三种备选理论，分别是温尼科特的“真实自我/虚假自我”理论、罗杰斯的“自我概念一致性”理论和埃里克森的“认同危机”理论(见附录 D)。这种多理论的提供有助于拓展研究者理论选择的视野，避免其陷入某一理论的思维定势。在获得备选理论后，大语言模型能进一步模拟不同理论视角下的解释路径，生成初步的竞争性假设。例如，模型可初步演绎出：基于温尼科特的理论，再嫁可能被解释为“虚假自我”对生存压力的适应，而离婚则是“真实自我”的回归。这种快速生成竞争性假设的能力，有助于激发研究者的比较与反思。

大语言模型跨理论检索的能力，在拓展视野的同时，也可能带来理论应用的“浅层化”与“去语境化”风险。大语言模型提供的理论，如温尼科特、罗杰斯或埃里克森的理论，是基于其知识图谱中与“自我”、“危机”、“婚姻”等关键词的语义关联度而被拼接推荐的。这种“拼图式”的推荐，缺乏对理论本身的历史脉络与文化语境的深度理解。例如，将源于西方个人主义文化背景的“真实自我”理论，直接用于分析宋代士人阶层女性李清照，需要评估其文化适配性。大语言模型可以模拟解释路径，但这种模拟往往是理论要点的机械推演，难以捕捉理论在面对具体、鲜活且浸透着特定历史文化的生命经验时所应有的谨慎、弹性与创造性。这可能导致研究陷入“理论套用”的陷阱：即从大语言模型给出的清单中选择一个“最相关”的理论，然后将传主生平作为例证填入。这与心理传记学致力于从独特的生命故事中孕育出新的概念理解(如“自我潮汐效应”)背道而驰。因此，研究者的角色应从“理论消费者”转变为“理论的批判性对话者与再创造者”，必须对大语言模型推荐的理论进行历史化、语境化的检视，并在理论与生命经验的碰撞中，发展出更具解释力的本土化概念框架。

3.5 高效识别理论盲点

心理传记学的生命力不仅在于运用已有理论解释传主，更在于发现已有理论的不足并发展新理论。如舒跃育(2009)从对诸葛亮的研究中抽象出了“二重形象”与“半杯水法则”；Xie 等人(2025)则在分析李清照时，提出了“自我潮汐效应”、“替代性自我实现”等新的概念。这些源于具体生命经验的创新，有助于推动中国本土心理学知识体系的构建。

在这一理论反思与建构环节，大语言模型能提供有力的过程性辅助。具体而言，当研究者选择某一理论后，大语言模型可以对其逻辑自洽性、与资料的契合度以及理论边界进行检验。以“李清照为何再嫁”这一问题为例，大语言模型首

先运用检索增强生成技术, 对温尼科特、罗杰斯、埃里克森的三种理论进行匹配度评估, 并建议优先采用温尼科特的虚假自我理论, 因其能同时解释“妥协再嫁”与“决然离婚”的矛盾行为。继而, 通过对抗样本检测技术(Adversarial sample detection), 模型识别出该理论的潜在失效场景, 并明确提示修正方向：需引入“宋代寡妇所面临的规范压力”这一文化语境。更具有反思性的是, 当模型被要求将其选出的“最佳”理论与研究者的原有理论进行比较时, 它得出后者更贴切的结论, 因为前者“偏重普适性心理冲突, 未充分纳入历史语境”(具体分析内容见附录 E)。

由此可见, 大语言模型确实能在整合多源信息的基础上, 帮助研究者检验理论、缩小选择范围并提示修正方向。但其“理论盲点检测”能力存在一定边界：大语言模型的检验, 是在既有的、已被编码的理论和数据中进行的。它擅长发现既有理论 A 在解释事实 B 时的不充分, 但它无法进行真正的理论创新——即提出一个全新的概念框架 C, 来更深刻地理解事实 B。例如, 大语言模型或许能指出用“中年危机”解释某个传主行为时的牵强之处, 但它无法像人类研究者那样, 从诸葛亮或李清照的具体生命经验中, 原创性地抽象出“半杯水法则”或“自我潮汐效应”这样的新概念。理论的真正创新, 往往源于研究者对异质个体生命不可化约的独特性的深刻敬畏与沉浸式理解, 并从中获得一种超越现有范畴的灵感。因此, 大语言模型在此环节更适合作为一个“假设生成器”和“盲点暴露者”, 迫使研究者更清晰、更严谨地陈述和捍卫自己的理论选择与创造。研究者的认识论优势, 最终必须体现在这种超越算法规则的创造性的理论建构能力上。

3.6　协同优化文本矛盾

心理传记文章的写作是一个循环往复的过程, 资料需要经常被重新访问和审查, 有时甚至需要补充额外的数据资料(Du Plessis, 2017)。随着人工智能能够将语言及其所包含的模糊性加以量化并赋予意义(刘忠波, 2021), 它可在语言、结构与论证逻辑等多个层面提供辅助, 从而提升写作效率。

在语言与结构层面, 模型可借助自动文摘技术快速梳理文献, 生成研究综述草稿, 或在保持原意的前提下, 对重复表述进行差异化润色, 优化行文流畅度。

在论证逻辑层面, 模型能够：1)可视化散落各处的证据网络, 辅助研究者检查论证断层或单一来源依赖; 2)通过整合文本、图像等多模态资料进行三角验证(Triangulation); 3)利用其数据偏见检测能力(蒋柯, 2024), 提示解释中可能隐含的“西方中心主义”等视角局限, 鼓励更具文化敏感性的分析。此外, 通过让模型扮演“反对者”角色, 可以检查文章前后论述是否存在矛盾, 或理论应用是否在传主生平的不同阶段存在不一致之处, 为其初步结论寻找反例或逻辑漏洞, 促使研究者完善论证。基于期刊审稿数据训练的模型, 还能模拟审稿评估(谭华 等, 2024), 指出文章不足并提供范式化的修改意见(Hosseini & Horbach, 2023)。

然而, 大语言模型的优化逻辑追求文本的一致性、流畅性与规范符合度, 这与心理传记学处理“悬疑性问题”的核心任务——保留并阐释生命固有的张力、矛盾与未解之谜——存在内在冲突。大语言模型识别出的“矛盾”, 可能是需要修补的逻辑漏洞, 但也可能正是传主生命中最核心、最值得深究的悬疑本身。如果盲目遵从大语言模型的优化建议, 反而可能抹平生命本身的复杂性, 生产出精致却平庸的叙事。

此外, 大语言模型基于主流学术范式训练的“润色”与“审稿”功能, 会施加无形的叙事同质化压力。对于李清照这类“异质”个体, 其生命故事可能需要突破常规的叙事结构与理论语言才能被恰当描述。大语言模型的优化可能在不经意间将这种独特的边缘经验纳入中心化的学术框架, 削弱其本应具有的理论颠覆性与思想冲击力(郑翔, 杨瑞仙, 2025)。因此, 在最终写作阶段, 研究者需要从“协同写作者”转变为最终叙事权威与审美判断的捍卫者。大语言模型“优化”的最高价值, 在于反向激发研究者更清醒地认识到自身主体性的不可替代。

综上所述, 大语言模型在心理传记研究中的应用, 远非简单的效率工具, 其介入引发了一场深刻的方法论与认识论反思。它不仅带来了资料真实性、算法偏见与语义理解等显性技术挑战, 更引发了研究者的主体性在技术逻辑的渗透下被稀释、外包乃至重构的危机。

4　大语言模型给心理传记学带来的挑战：主体性危机的多重表现

以上技术局限并非孤立存在, 而是构成了一

个环环相扣的挑战链条：输入资料所蕴含的数据污染和价值观偏见，催生了虚假或误导性内容，其根源又在于模型深层的语义理解局限无法真正把握历史语境与生命经验的复杂性。最终，这些多重挑战汇聚为一股合力，持续冲击着研究者在意义生成与叙事建构中的主导地位。下文将逐一剖析，这一主体性危机在心理传记研究的具体环节中，是如何通过数据污染、价值观偏见、虚假内容生成与语义局限等多重表现得以具体化与加深。

4.1 有偏见的输入：数据污染与内嵌的价值观偏见

大语言模型在处理资料过程中受数据影响较大(卢经纬 等, 2023; 赵京胜 等, 2019)，在心理传记研究中，这些作为“原料”的数据从源头便存在两类相互交织的缺陷——数据污染与价值观偏见。

首先，是数据污染。这主要是指信息本身的真实性、准确性与完整性缺失。网络信息本身的复杂性容易导致数据源污染，尤其是许多自媒体为吸引眼球、谋取私利，频繁制造和传播虚假信息(陈世华, 2022; 张媛 等, 2025)，进一步加剧了对大语言模型数据源的污染。在心理传记学中，这意味着关于传主的生平细节、关键事件叙述等可能混杂着大量虚构情节。例如，网络上广为流传的某些历史人物的“名言”或“秘闻”，往往缺乏原始出处却叙事生动，极易被模型视为有效数据。当这些被污染的数据被整合进一份看似客观的资料清单时，就容易使得后续的分析失真。

其次，是更深层、更隐蔽的价值观偏见。大语言模型的技术逻辑不可避免地承载着特定的社会价值观(张秀华, 郭静茹, 2025)，其训练依赖的海量数据多为主流文化筛选后的产物，非主流个体的生命经验往往被排除在外。在训练过程中，如果数据集中包含了性别歧视、种族歧视等不良价值观，那么模型在生成内容时也可能表现出这些偏见。例如，由于训练数据中英语资料占比较大，导致 DeepSeek 等大语言模型在选择传主时，推荐的外国名人往往多于中国名人(王均松, 2025; 张晨晓, 2025; 周利生, 刘芳华, 2025)。这种数据偏向通过统计学构建的“因果链”，被包装成逻辑自洽的分析结论，从而提升了大语言模型的“科学性” (李佳泓 等, 2024)，但其背后依然隐藏着逻辑错误、矛盾及偏见(汪凡淙 等, 2025)。如“GPT-3 模型在文本补全任务中会产生针对女性的性暴力内容” (Wyer & Black, 2025)。

更为关键的是，这两种缺陷在实践中相互强化，形成叠加效应。一份被广泛传播的虚假史料，其内容往往反映了某种特定的社会偏见；反之，一种占主导地位的价值观偏见，又会决定哪些内容被优先保存、数字化和反复引用。当大语言模型基于这样的“原料”进行学习与推理时，其生成的认知框架与内容输出，从源头便偏离了全面与客观的轨道，为其生成结果的虚假性埋下了伏笔(Helberger & Diakopoulos, 2023)。

4.2 有偏差的输出：虚假内容与偏见叙事的生成

在数据源被污染的情况下，大语言模型由于缺乏对资料真实性的鉴别能力(何宇华，李霞，2024)，导致虚假错误资料被纳入分析框架，从而“被动”生成错误结论。例如，在分析某个历史人物的传记时，如果数据源中存在关于该人物的虚假信息，大语言模型在生成相关内容时可能会将其错误地纳入其中，导致研究结果失真。在心理传记研究的具体实践中，这种偏差主要表现为两种：虚假内容的生成与偏见性叙事框架的强化。它们共同威胁着研究结论的可靠性，并悄然将研究者的分析引向歧途。

首先，是虚假内容的生成。大语言模型不具备事实判断能力，这种技术逻辑决定了其输出的内容并非基于对传主真实意图或历史背景的理解，而是对已有语料库的统计性重组和逻辑推论。当面对史料空白、记载矛盾或模糊之处时，模型并非像人类研究者那样承认“未知”并寻求新证据，而是倾向于基于既有文本模式进行“合理”的补全或推断。即使输入随机内容，它也会机械地“强行给出答案”，一旦被指出错误便立即道歉。这种“强行给出的答案”往往会对不熟悉该领域的人造成误导(李佳泓 等, 2024; 维克托莉亚·克雷齐格，邓环宇, 2025)。

其次，是对偏见性叙事框架的强化，这源于模型对数据中主流模式的放大。大语言模型倾向于生成最符合其训练数据中高频、主流模式的解释和叙述(Barnes et al., 2024)。在心理传记分析中，这意味着模型会不自觉地调用并加固现有的、可能充满偏见的解释模板。例如，在分析李清照时，大语言模型更可能从其数据中关联并强调“情感创伤驱动创作”、“性格叛逆源于家庭缺失”等常见但可能简化的心理叙事，而非去构建一个更复

杂、更超越时代性别框架的、基于智性追求或主体性抗争的解释体系。这种输出便是一种解释性偏见, 它使得研究者的理论视角被无形地窄化, 消解了异质个体本该具有的理论冲击力。

因此, 大语言模型的输出并非中性工具的结果。它是在有缺陷的输入和自身技术逻辑共同作用下的产物。研究者如果缺乏高度的警惕, 便可能将这个由算法重构的、可能既包含事实谬误又充满叙事偏见的“模型世界”当作传主的内心世界, 从而冲击和削弱了研究者在研究中的主体地位。而这些问题的根源在于大语言模型技术中无法逾越的鸿沟——对语义理解的局限。

4.3　根本的鸿沟: 语义理解局限与意义生成缺失

无论是有偏见的输入还是有偏差的输出, 其最深层的根源在于大语言模型固有的语义理解局限。

首先, 大语言模型缺乏对具体语境与语义的理解能力。大语言模型通过识别文本中具有特定意义的词语, 获取实体之间的语义关联(赵京胜等, 2019), 仅涉及浅层语义分析, 缺乏理解深层语义的能力(车万翔 等, 2023)。当面对模糊或不明确的语境时, 大语言模型只能根据已有资料生成多种解释(Zhang et al., 2023), 无法像人类研究者那样, 借助文化直觉、历史共情和生命经验进行深度解读(Kiryakova & Angelova, 2023)。例如, 有关中国古代历史人物的记载往往是以文言文的形式保存下来的。这些史料在语法、词汇和表达方式上与现代汉语存在诸多不同, 需要依托古代汉语语法知识与历史文化语境方能准确解读(张秋玲, 2010)。现有的大语言模型虽可通过大规模数据训练实现文本处理(傅安国, 岳童, 2024), 但仍存在因中文训练数据不足导致的文化适应性薄弱, 以及因缺乏文言文深度训练而导致的解读准确性低下等问题(张华平 等, 2023; Li et al., 2024), 如“GPT-4 在对文言文语义解读的过程中会忽视宾语前置的问题, 出现对文本语义的错误判断”(高承海 等, 2025), 导致其无法准确把握语句表达的深层含义, 也无法像人类一样体会到语言背后的文化意蕴(Rimban, 2023)。

其次, 模型不具备真正的“意义生成”能力。心理传记学的目的, 是整合碎片化的事实, 运用理论, 编织出一个有说服力、有深度、能照亮独特心灵的故事。这是一个创造性的意义生成过程, 依赖于研究者的理论直觉、人文共情、叙事智慧以及对生命复杂性的敬畏。大语言模型可以拼接术语、概括生平、模仿学术表述, 但其物理式的认知方式无法理解人类独特的精神现象与意识体验(李微, 张荣军, 2022)。“当传主的悬疑性问题能够让研究者产生惊奇感和震撼感, 从而激发研究灵感时”(舒跃育, 2024), 人工智能仅能生成基于概率的“合理却平庸”的答案, 无法在深层语义层面实现与传主的对话, 进一步放大虚假性内容, “难以规避文化误读引发的伦理风险”(郑剑虹, 2024)。

因此, 语义理解局限是其技术缺陷背后的根本原因。它意味着, 即便输入数据相对干净, 输出形式看似规范, 模型最终提供的也只是一个由符号关联编织的、缺乏生命温度的“语义表象”。研究者如果过度依赖其结果, 便会与传主真实、鲜活、充满矛盾与深意的生命经验失之交臂。

4.4　主体的退场: 研究者在技术流程中责任主体的让渡

心理传记学作为探索人性边界的学科, 其核心在于通过对传主资料的创造性诠释, 发掘人性边缘处的“真、善、美”, 这一过程本质上要求研究者保持高度的理论敏感性(Elms, 2005, pp. 84−95)。当数据源的污染、输出内容的偏差以及深层的语义隔阂, 共同构成了一套环环相扣的技术逻辑链条时, 很可能让研究者悄然让渡了其作为诠释主体的核心地位(张灿, 郭海静, 2023), 冲击研究者在心理传记研究中的主体性地位。

首先, 当大语言模型能够快速完成资料筛选、编码甚至生成理论草稿时, 研究者可能陷入对技术效率的迷恋, 将主要精力投入如何优化提示词、调试参数(例如前文中许多研究者致力于提示词优化、分析框架适配等研究), 而非沉浸于对传主生命本身的沉思与共情。使得心理传记学特有的“悬疑性解释”沦为程序化的算法推导过程, 其生成的“创造性”成果本质上只是既有数据的排列组合(AlZu'bi et al., 2024), 缺乏深刻的思想性与灵活性(Nazir & Wang, 2023), 抹杀了研究主体的创造性(张秀华, 郭静茹, 2025)。

其次, 当人工智能的使用变得不可避免时, 研究者可能倾向于选择那些更容易被模型解释的传主和问题, 避开那些真正复杂、模糊但更具价值的“硬骨头”。使得研究过程从“以问题为中心”

异化为“以技术和方法为中心”, 不仅削弱了对文本的诠释深度(蒋柯, 2024), 更可能使研究者全盘接受人工智能基于数据重组构建的“伪创新”理论(郑剑虹, 2024), 阻碍真正具有创新性的理论的产生和发展, 缩减研究者的创造性思维。

最后, 面对模型输出的结构清晰、表述专业的分析报告, 研究者可能不自觉地将其权威化, 压抑自身的批判性质疑和不同寻常的解读灵感。长此以往, 研究者的理论创造力和学术自信可能退化, 最终默认算法推导出的“最优解”, 使心理传记学丧失其作为人文学科的批判性与创造性灵魂。归根结底, 即便是最先进的技术辅助系统, 其认知价值仍完全依附于研究者的创造性思维(汪凡淙 等, 2025)。在技术工具日益强大的当下, 研究者需要警惕“分析外包”的诱惑(何吴明, 郑剑虹, 2019), 保持自身的敏感性和创造性, 将人工智能更多作为一种辅助性工具(Anantrasirichai & Bull, 2022), 坚守研究者作为理论创造主体的核心地位。

综上所述, 大语言模型带来的主体性危机, 是一个从“外围辅助”向“核心领域”逐步渗透的慢性过程。它通过提供看似高效、客观、全面的服务, 使研究者在享受便利的同时, 不知不觉中交出了发现问题、评估证据、创造理论等关键权能。因此, 研究者必须在每一个环节保持清醒的“认识论警惕”, 明确划清技术能力的边界, 并承担起技术无法替代的责任。

5 技术发展引发的心理传记学研究中主体性危机的应对

然而, 这些危机并非既定结局。面对技术逻辑的渗透, 研究者必须从被动适应转向主动建构, 通过技术治理与流程重塑, 在人工智能时代捍卫自身在意义诠释与理论创造中的核心地位。下文将从数据基础、伦理规范与方法流程三个层面, 探讨应对主体性危机的可能路径。

5.1 加强本土文化数据库的建设以应对数据偏见与跨文化解读

大语言模型训练依赖的海量数据多为主流文化筛选后的产物, 其技术逻辑与生成内容不可避免地带有特定的社会文化价值观, 容易引发和加剧人类偏见(李佳泓 等, 2024; 张秀华, 郭静茹, 2025; Friedrich et al., 2025; Zhang et al., 2023)。这些问题的根源在于大语言模型训练与检索的数据源的文化片面性和可信度缺失。因此, 重构技术基础的首要任务是确保数据的准确性与全面性, 在心理传记学领域则具体表现为开发基于本土文化的语料库, 以此作为大语言模型的数据分析库。

以我国为例, 中国作为世界上少有的拥有五千年未曾中断的独特文化的国家, 形成了诸多有助于理解人性的理论, 且对重要历史人物有着标准的史料记载, 是未来心理传记学发展的主要阵地(舒跃育, 2018b)。未来可基于国内丰富的史料数据, 开展史料的数字化工作, 形成专业的本土文化数据库, 作为对中国古代历史人物进行心理传记学研究的可靠、便捷的资料来源。如数字人文领域中的古籍大语言模型“荀子”的开发(南京农业大学新闻网, 2023), 显示出当前已存在实现的技术可能。

除了心理传记学领域, 未来还可以根据研究需要, 联合不同学科学者构建针对不同社会文化的多样化的数据库, 再基于该数据库训练特定的大语言模型, 形成“构建本土文化资料库与大语言模型, 助力于相关研究, 再以相关研究成果修正和调整模型的文化准确性”的循环研究模式(见图 1), 确保特定大语言模型的文化敏感性和专业性。这样既可以保证资料的可靠性, 又能够避免异文化对本土文化的误读。

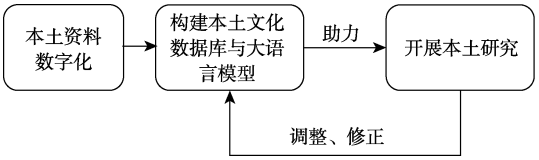

图 1 本土文化数据库驱动的大语言模型与研究循环

在通过构建本土文化数据库以避免大语言模型及其训练数据可能带有的偏见之后, 如何确保研究者在使用这些数据与 AI 工具时, 能够明确并履行其主体责任, 便成为下一个关键议题。这需要建立清晰的伦理原则框架。

5.2 完善大语言模型在心理传记学中应用的伦理原则以明确研究者的主体责任

自 2023 年起, 论文中使用人工智能的研究比例呈快速增长趋势, 但仅有 0.1%的作者在正文中明确披露曾使用人工智能进行写作辅助(He & Bu, 2026)。在人工智能的使用成为大趋势的情况下, 如何规避大语言模型的加入给心理传记学带来更

多的伦理问题? 学术共同体需要构建并完善人工智能在心理传记学中应用的伦理原则, 明确研究者在使用人工智能时的主体责任与行为边界, 并将系统的伦理审查贯穿研究始终。基于前文研究, 本文提出以下几个原则可供参考:

人类主体性与责任归属原则。在心理传记的研究中, 大语言模型应被严格限定为一种“辅助工具”, 研究者必须始终保持自己在研究中的主体性。在研究过程中, 所有使用 AI 生成的内容, 研究者都必须进行严格的核查、批判与整合。在发表研究成果时, 研究者必须明确披露 AI 在研究中的参与程度和具体用途, 最大限度地保持 AI 使用的透明度。在责任归属方面, AI 不应也不能成为独立的作者, 研究者始终是研究和论文等学术成果的第一责任人。

公平性与偏见修正原则。在使用大语言模型进行心理传记研究的过程中, 研究者要始终保持自身的批判与反思能力, 主动识别并修正算法偏见。包括: 审慎选择和分析训练数据可能存在的缺陷; 在利用大语言模型进行文本分析时, 采用多种理论视角进行交叉验证; 将研究结果放回到传主所处的具体历史与社会文化语境的理解中, 避免被模型的“平均化”输出所误导。

隐私保护与数据安全强化原则。心理传记学研究往往涉及大量与传主相关的私人资料(如私人信件、日记、著作等)。这些资料的使用应严格限制在本土语料库中, 在输入方面, 要对资料进行多重清理, 确保消除可能包含的污染数据。在输出环节, 严禁将可识别传主身份的敏感原始数据直接输入公共大语言模型, 以防数据被用于模型训练而导致信息泄露。

风险预防与动态治理原则。鉴于大语言模型可能被“指令攻击”诱导生成有害内容(陈晋音 等, 2025), 以及心理传记学研究本身可能对传主声誉、其后人或相关群体造成的潜在影响, 必须建立“风险——价值”评估机制(舒跃育 等, 2021; Ponterotto, 2018)。在研究前, 研究者需说明研究目的、研究价值及潜在影响, 供伦理审查委员会评估该研究的必要性。在研究成果发表时, 期刊需要考察其研究逻辑与结果是真实或合理的。若研究发表后引发争议, 应提供伦理申诉渠道, 并进行公开回应与处理。同时, 鉴于大语言模型技术发展的快速性, 伦理规范也需定期复审与更新, 以应对可能出现的各种新的问题。

上述伦理原则明确了研究者的责任边界, 而要将其落到实处, 则需要在具体的研究流程设计中, 确保人类智能始终占据主导地位, 实现有效的人机协同。

5.3 构建研究者主导的人机协同研究流程以促进人工智能与人类智能的互补

为应对前文所述挑战, 确保研究者的主体性, 本文提出了以研究者为主导的人机协同研究流程。该流程旨在将人类智能的诠释深度、理论直觉与伦理判断, 与人工智能的信息处理、模式识别与计算效率相结合, 实现优势互补。下文将以“悬疑性问题”分析模式为例, 阐述这一协同流程的具体环节, 并着重说明在各环节中研究者不可替代的核心作用。

在前文基础上, 本文绘制了人机协同研究流程图(见图 2)。该流程并非线性, 而是一个循环迭代、人机交互的有机整体。在整个研究过程中, 始终需要研究者保持足够的伦理考量, 并对生成内容负责, 确保其符合学术伦理规范。具体可分为以下几个步骤:

(1)研究者确立研究核心: 研究者基于学术问题意识, 选定传主并界定理论视角。此步骤依赖研究者独特的研究视角。

(2)大语言模型对海量文献进行初步检索与收集: 研究者输入指令后, 大语言模型可运用检索增强生成技术, 从指定的专业数据库、数字化档案及经过审核的网络资源中, 快速抓取与传主相关的海量文献(如生平记载、著作、书信、历史背景资料等), 并生成初步的文献清单与摘要。

(3)研究者对资料进行人工核查: 研究者对大语言模型收集的资料库进行批判性审阅, 核实关键信息, 剔除错误内容。此步骤是确保研究内容真实性的关键, 大语言模型无法替代研究者对史料真伪与价值做出专业判断。

(4)研究者定义分析维度并微调模型: 研究者根据研究目的, 将心理传记学已有的资料筛选理论(如“原型情景理论”、“凸显性指标”、“成长性关键因素”等), 转化为具体、可操作的分析维度和标注规则。例如, 根据“原型情景理论”, 研究者可设定“识别资料中反复出现的、具有情感象征意义的场景或意象(如‘黄昏独酌’、‘战乱南渡’)”作为核心分析维度, 并提供标注示例。研究者进而通过领

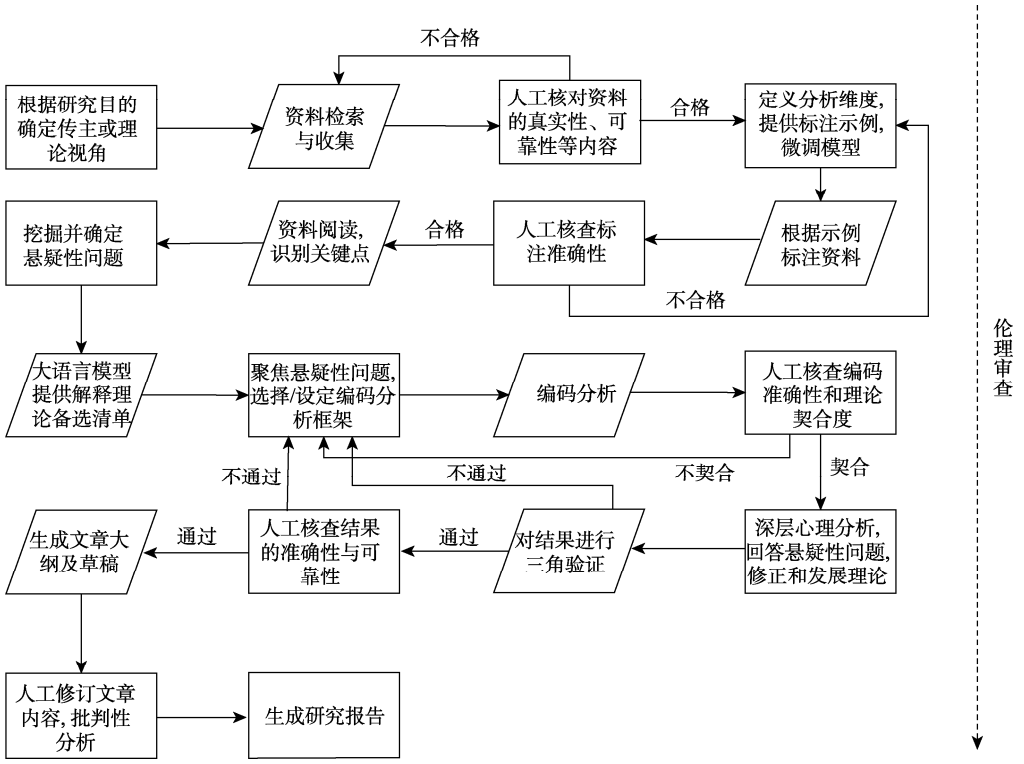

图 2 人机协同研究流程图

注：矩形方框代表以研究者为主的工作任务，菱形方框代表大语言模型可以提高效率的工作任务。本流程图及框架为基于现阶段探索形成的初步构想，其有效性与普适性尚需后续研究实践进一步检验和细化。

域适应性微调(Domain Adaptation Fine-tuning)或构建专业提示词工程(Prompt Engineering)，将此理论化的维度体系“注入”大语言模型，使其后续的识别与编码工作严格遵循研究者预设的分析逻辑，而非进行通用语义处理。

(5)大语言模型辅助深度阅读与初步模式识别：利用微调后的模型，对已审核的资料进行批量处理。大语言模型可运用自然语言处理技术(如命名实体识别、情感分析、主题建模等)，识别文本中的关键模式(如情感演变轨迹、主题网络、人物关系动态)，生成初步的资料分析报告。

(6)研究者确定悬疑性问题，设定分析框架：研究者基于自身对传主的初步理解与共鸣，结合大语言模型提供的分析报告及深刻的社会文化背景知识，进行“主观悬疑与客观清单的共鸣验证”。通过这一对话过程，研究者最终提炼并确定核心“悬疑性问题”。随后，利用大语言模型跨领域知识图谱的优势，针对该具体问题快速生成一个多元的“解释理论备选清单”(例如，对于“李清照再嫁的心理动因”，模型可能推荐温尼科特的“真实自我”理论、埃里克森的“自我认同”理论或关于“创伤后应对”的认知行为理论等)。研究者可将此清单作为启发式参考，结合自身理论素养与研究发现，作出主导性的初步理论选择，并据此设定具体的编码分析框架。

(7)人工核查编码准确性和理论契合性：此后的编码与分析过程，是一个“理论——资料”动态适配的检验过程。研究者需要基于自身对传主资料的理解，批判性地评估模型生成的编码内容的准确性与合理性。若恰当则进入下一环节，不恰当则返回步骤(6)调整编码分析框架，并重新开始编码。这个过程可能循环数次，直至找到或发展出一个具有足够解释力的框架。

(8)深层心理分析、悬疑性问题解答与理论发展：当研究者确定了最终解释的理论框架后，需要再回到传主资料中，评估其与传主生命经验的契合度，进行创造性整合与理论修正，形成对悬疑性问题的初步解释。同时，对传主资料分析得

出的结果与已有理论之间的差异之处, 便很可能是该理论得以发展的关键点。研究者可基于此进一步发展或厘清该理论的应用范围。

(9)人机协同三角验证：研究者与大语言模型共同对初步结论进行验证。利用大语言模型通过多源验证法(如对比传主自述、他人评价、历史记录)或多模型对比法(使用不同分析模型复现结论), 对结果进行交叉检验。研究者则综合大语言模型的验证结果与自身的学术判断, 评估结论的可靠性。

(10)大语言模型生成报告初稿, 研究者进行最终修订与升华：在结论通过验证后, 可让大语言模型根据既定的论述逻辑和证据, 生成研究报告或论文的大纲及部分草稿。研究者必须对全文进行人工修订、深化与升华, 注入其独特的叙事风格、理论创见与人文关怀, 确保成果是研究者智力劳动的体现, 而不仅仅是大语言模型的格式化输出。

总体来说, 以上流程是一个“研究者主导、技术赋能”的循环迭代模型。在每一个环节, 研究者的核心作用都不可或缺：提出问题、设定规则与框架、做出关键判断、进行深度诠释与最终负责。大语言模型的作用则体现在：处理海量信息计算、提供多角度模式识别、辅助逻辑验证与草稿生成。这一设计流程, 旨在通过明确的人机分工与紧密交互, 将人工智能与人类智能相结合, 最终实现“1 + 1 > 2”的协同效应, 在提升研究效率的同时, 捍卫研究者在意义生成与理论创造中的主体地位。

最后, 以上人机协同研究流程是基于本文中采用的“悬疑性”问题分析模式而提供的一个参考框架和初步探索。在具体研究过程中, 研究者应发挥自身的创造性思维、批判性反思和伦理敏感性, 始终保持对研究问题、理论视角、解释框架的主导权, 根据研究的不同目的、不同阶段, 对人工智能的介入程度与时机进行灵活调整, 以进一步完善人机协同研究框架及其实践路径。

6　结语

人工智能的出现为心理传记研究带来了前所未有的机遇与挑战, 本文从理论层面探讨了大语言模型在心理传记学研究中的技术优势与潜在挑战, 提出了本土化语料库驱动下的大语言模型循环研究模式、大语言模型在心理传记学中应用的伦理准则, 以及研究者主导下的人机协同研究框架, 深化了对大语言模型在人文社会科学领域应用的理论思考, 肯定了人是研究中的主体, 人工智能只是辅助工具的角色定位, 为学科研究范式的革新提供了理论支撑与方向指引。但与此同时, 本研究结果仅基于大语言模型 DeepSeek 与 ima-AI 智能体所得出的个案, 不宜简单、直接外推至其他大语言模型或所有心理传记研究情境中, 而需要在具体的研究过程中, 结合研究内容与应用模型进行实际的操作验证。在理论创新与深度诠释层面, 当前的大语言模型也仅适合作为一个“假设生成器”和“盲点暴露者”, 辅助拓宽研究者的视野, 而非“理论的最终裁决者”。

归根结底, 人工智能的介入并未改变心理传记学的核心价值：通过对异质个体生命的深度诠释, 展示人性的复杂性与独特性。当前大语言模型的优势在于效率提升与模式识别, 其不可逾越的局限在于意义生成与理论创造。本文提出的本土化语料库、伦理准则和人机协同框架, 仅仅是对人工智能时代下心理传记学方法论探索的开端, 真正决定心理传记学未来的, 始终是研究者的理论敏感性与人文关怀。

参考文献

曹家齐. (2022). 史观与史实——评《宋高宗新论》. *中华文史论丛*, *63*(2), 389−395.

车万翔, 窦志成, 冯岩松, 桂韬, 韩先培, 户保田, ... 赵妍妍. (2023). 大模型时代的自然语言处理: 挑战、机遇与发展. *中国科学: 信息科学, 53*(9), 1645−1687.

陈晋音, 席昌坤, 郑海斌, 高铭, 张甜馨. (2025). 多模态大语言模型的安全性研究综述. *计算机科学, 52*(7), 315−341.

陈世华. (2022). 以假乱真与去伪存真: 自媒体欺骗行为的表征及其治理. *学习与实践, 39*(6), 132−140.

陈雪仪. (2021). *事业与爱情: 林徽因重大人生决策的心理传记学分析* [硕士学位论文]. 西北师范大学.

丁邱, 迟海洋, 严馨, 徐广义, 邓忠莹. (2023). 基于 Transformer 模型的问句语义相似度计算. *计算机工程与设计, 44*(3), 887−893.

傅安国, 岳童. (2024-10-16). ChatGPT 与质性研究的融合. *中国社会科学报*. 第 A06 版. https://epaper.csstoday.cn/epaper/read.do?m=i&iid=6936&eid=49976&sid=231425

高承海, 党宝宝, 王冰洁, 吴胜涛. (2025). 人工智能的语言优势和不足: 基于大语言模型与真实学生语文能力的比较. *心理学报, 57*(6), 947−966.

何吴明, 郑剑虹. (2019). 心理学质性研究: 历史、现状和展望. *心理科学, 42*(4), 1017−1023.

何宇华, 李霞. (2024). 生成式人工智能虚假信息治理的新挑战及应对策略——基于敏捷治理的视角. *治理研究, 40*(4), 142−156+160.

何忠礼. (2021). *宋高宗新论*. 上海: 上海古籍出版社.
黄希庭. (2021). *心理学研究方法*. 重庆: 西南大学出版社.
蒋怀滨, 张斌. (2019). *心理学研究方法*. 厦门: 厦门大学出版社.
蒋柯. (2024-07-03). 大语言模型与质性研究技术的耦合. *中国社会科学报*. 第 A05 版. https://epaper.csstoday.cn/epaper/read.do?m=i&iid=6859&eid=49246&sid=228027
李佳泓, 黄嘉, 古炬贤. (2024). 论自然语言处理中的意识形态安全问题——以 ChatGPT 为例. *情报杂志*, *43*(8), 1−9.
李微, 张荣军. (2022). 价值理性视域下的人工智能及其伦理边界. *贵州社会科学*, *43*(5), 13−20.
刘忠波. (2021). 人工智能写作意味着什么? ——人工智能时代的写作主体问题. *南京师范大学文学院学报, 23*(4), 2−9.
卢经纬, 郭超, 戴星原, 缪青海, 王兴霞, 杨静, 王飞跃. (2023). 问答 ChatGPT 之后: 超大预训练模型的机遇和挑战. *自动化学报, 49*(4), 705−717.
麦克亚当斯. (2024). *人格心理学: 人的科学导论* (第 5 版) (郭永玉 主译). 上海: 上海教育出版社. (原著出版于 2009 年)
南京农业大学新闻网. (2023-12-07). *首创! 我校研发的全国首个古籍大语言模型正式发布*. 南京农业大学. 2025-06-01 取自 https://news.njau.edu.cn/2023/1206/c18a127132/pagem.htm
潘晓英, 陈柳, 余慧敏, 赵逸喆, 肖康泞. (2020). 主题爬虫技术研究综述. *计算机应用研究, 37*(4), 961−965+972.
钱慎一, 付博文, 李代祎, 梁瑶瑶. (2025). 面向知识图谱的问答技术研究综述. *计算机工程与应用*, *62*(23), 1−25.
秦小林, 古徐, 李弟诚, 徐海文. (2025). 大语言模型综述与展望. *计算机应用, 45*(3), 685−696.
任欢, 王旭光. (2021). 注意力机制综述. *计算机应用, 41*(S1), 1−6.
尚钰, 刘锟, 韩霄龙. (2024). 基于大语言模型智能体的跨域立场检测. *信息安全与通信保密*, *46*(8), 62−71.
舒尔茨. (2011). *心理传记学手册* (郑剑虹 等 译). 广州: 暨南大学出版社. (原著出版于 2005 年)
舒跃育. (2009). *历史人物之二重形象研究——以诸葛亮的心理传记分析为例* [硕士学位论文]. 西北师范大学.
舒跃育. (2018a). *天命可违——诸葛亮行为决策的心理传记学分析*. 北京: 清华大学出版社.
舒跃育. (2018b). 心理传记学的历史与展望. *西北师大学报(社会科学版), 55*(5), 102−109.
舒跃育. (2024-07-03). 人工智能视域的质性心理学研究. *中国社会科学报*. 第 A05 版. https://epaper.csstoday.cn/epaper/read.do?m=i&iid=6859&eid=49246&sid=228029
舒跃育, 王雪婵, 郑剑虹, 石莹波, 汪李玲, 袁彦. (2021). 作为分支学科的心理传记学: 当前困境与应对. *心理科学, 44*(3), 761−767.
谭华, 习琲, 王晓丹, 朱天潇. (2024). 人工智能技术赋能高校科技期刊高质量发展的应用研究. *漯河职业技术学院学报, 23*(6), 87−91.
汪凡淙, 汤筱玓, 余胜泉. (2025). 基于生成式人工智能的认知外包: 交互行为模式与认知结构特征分析. *心理学报, 57*(6), 967−986.
王均松. (2025-12-10). 警惕大语言模型的语用偏见. *中国社会科学报*. 第 A09 版. https://epaper.csstoday.cn/epaper/read.do?m=i&iid=7269&eid=53160&sid=247553
王泽钰. (2022). *女性主体性: 武则天与慈禧的比较心理传记学研究* [硕士学位论文]. 西北师范大学.
王曾瑜. (2016). *宋高宗传*. 北京: 中国书籍出版社.
维克托莉亚·克雷齐格, 邓环宇. (2025). 针对人工智能生成影像侵权的保护. *南大法学*, *32*(2), 168−184.
谢霞. (2022). *女性真实自我的寻求: 以李清照的心理传记学分析为例* [硕士学位论文]. 西北师范大学.
颜文靖. (2024-07-03). 大语言模型开启质性研究编码分析新时代. *中国社会科学报*. 第 A05 版. https://epaper.csstoday.cn/epaper/read.do?m=i&iid=6859&eid=49246&sid=228028
叶浩生. (2009). 西方心理学中的质化研究思潮. *社会科学, 31*(11), 113−118.
张灿, 郭海静. (2023). 数字时代数字技术崇拜的机制与实质. *重庆邮电大学学报(社会科学版), 35*(2), 93−102.
张晨晓. (2025). 突围与博弈: 从 DeepSeek 现象看大语言模型时代文化自信的建构. *郑州铁路职业技术学院学报, 37*(4), 103−107.
张华平, 李林翰, 李春锦. (2023). ChatGPT 中文性能测评与风险应对. *数据分析与知识发现*, *7*(3), 16−25.
张培晶, 宋蕾. (2012). 基于 LDA 的微博文本主题建模方法研究述评. *图书情报工作, 56*(24), 120−126.
张秋玲. (2010). 文言文“浅易”的语词特征研究——以百年来初中教科书中的文言选篇为研究对象. *语言文字应用, 19*(3), 115−122.
张秀华, 郭静茹. (2025). 人工智能语境下文化主体性危机及应对. *自然辩证法研究, 41*(3), 9−17.
张媛, 王昕然, 何云峰. (2025). 互联网短视频内容的生产、逻辑及治理. *融媒*, *2*(4), 15−22.
赵京胜, 宋梦雪, 高祥. (2019). 自然语言处理发展及应用综述. *信息技术与信息化*, *50*(7), 142−145.
郑剑虹. (2013). 中国大陆的心理传记学研究及其质量结合模式. *生命叙事与心理传记学, 1*(1), 41−61.
郑剑虹. (2024-10-16). 人工智能技术在质性研究中的应用. *中国社会科学报*. 第 A06 版. https://epaper.csstoday.cn/epaper/read.do?m=i&iid=6936&eid=49976&sid=231427
郑翔, 杨瑞仙. (2025). AI 可以代替学者审稿吗?——基于大语言模型的审稿意见生成与对比研究. *情报杂志, 44*(10), 1−9.
周利生, 刘芳华. (2025). DeepSeek 类生成式人工智能嵌入意识形态治理的价值、风险与应对. *南昌大学学报(人文社会科学版), 56*(2), 40−49.
周明洁, 张建新. (2008). 心理学研究方法中“质”与“量”的整合. *心理科学进展, 16*(1), 163−168.
AlZu'bi, S., Mughaid, A., Quiam, F., & Hendawi, S. (2024). Exploring the capabilities and limitations of ChatGPT and alternative big language models. *Artificial Intelligence and Applications, 2*(3), 28−37.
Anakok, I., Katz, A., Chew, K. J., & Matusovich, H. (2025). Leveraging generative text models and natural language

processing to perform traditional thematic data analysis. *International Journal of Qualitative Methods, 24,* 1−13. https://doi.org/10.1177/16094069251338898

Anantrasirichai, N., & Bull, D. (2022). Artificial intelligence in the creative industries: A review. *Artificial Intelligence Review, 55*(1), 589−656.

Barnes, A. J., Zhang, Y., & Valenzuela, A. (2024). AI and culture: Culturally dependent responses to AI systems. *Current Opinion in Psychology*, *58*(4), 101838. https://doi.org/10.1016/j.copsyc.2024.101838

Barrera, B., Feliu, A., Espina, C., Brand, T., Zeeb, H., & Ahmed, F. (2025). Leveraging AI to enhance qualitative research: Experiences and recommendations from case studies in cancer prevention literacy across the European Union. *International Journal of Qualitative Methods, 24*, 1−10. https://doi.org/10.1177/16094069251365766

Brondani, M., Alves, C., Ribeiro, C., Braga, M. M., Garcia, R. C. M., Ardenghi, T., & Pattanaporn, K. (2024). Artificial intelligence, ChatGPT, and dental education: Implications for reflective assignments and qualitative research. *Journal of Dental Education*, *88*(12), 1671−1680.

Chatzichristos, G. (2025). Qualitative research in the era of AI: A return to positivism or a new paradigm? *International Journal of Qualitative Methods, 24*, 1−12. https://doi.org/10.1177/16094069251337583

Chen, T.-J. (2023). ChatGPT and other artificial intelligence applications speed up scientific writing. *Journal of the Chinese Medical Association, 86*(4), 351−353.

Chopra, F., & Haaland, I. (2023). *Conducting qualitative interviews with AI* (CEBI working paper series No. 23-06). Center for Economic Behavior and Inequality (CEBI), University of Copenhagen. https://ideas.repec.org/p/cep/cebiwp/23-06.html

De Paoli, S. (2024). Performing an inductive thematic analysis of semi-structured interviews with a large language model: An exploration and provocation on the limits of the approach. *Social Science Computer Review, 42*(4), 997−1019.

Du Plessis, C. (2017). The method of psychobiography: Presenting a step-wise approach. *Qualitative Research in Psychology, 14*(2), 216−237.

Elms, A. C. (2005). If the glove fits: The art of theoretical choice in psychobiography. In W. T. Schultz (Ed.), *Handbook of psychobiography* (pp. 84−95). Oxford University Press.

Friedrich, F., Brack, M., Struppek, L., Hintersdorf, D., Schramowski, P., Luccioni, S., & Kersting, K. (2025). Auditing and instructing text-to-image generation models on fairness. *AI and Ethics, 5*(3), 2103−2123.

Gibson, A. F., & Beattie, A. (2024). More or less than human? Evaluating the role of AI-as-participant in online qualitative research. *Qualitative Research in Psychology, 21*(2), 175−199.

Gur, T., & Maaravi, Y. (2025). The algorithm of friendship: Literature review and integrative model of relationships between humans and artificial intelligence (AI). *Behaviour & Information Technology, 44*(14), 3446−3466.

Hamilton, L., Elliott, D., Quick, A., Smith, S., & Choplin, V. (2023). Exploring the use of AI in qualitative analysis: A comparative study of guaranteed income data. *International Journal of Qualitative Methods, 22*, 1−13. https://doi.org/10.1177/16094069231201504

He, Y., & Bu, Y. (2026). Academic journals' AI policies fail to curb the surge in AI-assisted academic writing. *Proceedings of the National Academy of Sciences, 123*(9), e2526734123.

Helberger, N., & Diakopoulos, N. (2023). ChatGPT and the AI Act. *Internet Policy Review, 12*(1), 1−6.

Hitch, D. (2024). Artificial intelligence augmented qualitative analysis: The way of the future? *Qualitative Health Research, 34*(7), 595−606.

Hosseini, M., & Horbach, S. P. (2023). Fighting reviewer fatigue or amplifying bias? Considerations and recommendations for use of ChatGPT and other large language models in scholarly peer review. *Research Integrity and Peer Review, 8*(1), 4.

Kiryakova, G., & Angelova, N. (2023). ChatGPT—A challenging tool for the university professors in their teaching practice. *Education Sciences, 13*(10), 1056.

Li, Z., Qiu, W., Ma, P., Li, Y., Li, Y., He, S., ... Gu, W. (2024). An empirical study on large language models in accuracy and robustness under Chinese industrial scenarios. *arXiv*. https://doi.org/10.48550/arXiv.2402.01723

Liang, P. P., Wu, C., Morency, L.-P., & Salakhutdinov, R. (2021). Towards understanding and mitigating social biases in language models. In M. Meila & T. Zhang (Eds.), *Proceedings of the 38th International Conference on Machine Learning* (Vol. 139, pp. 6565−6576). PMLR.

Marshall, D. T., & Naff, D. B. (2024). The ethics of using artificial intelligence in qualitative research. *Journal of Empirical Research on Human Research Ethics*, *19*(3), 92−102.

Mathis, W. S., Zhao, S., Pratt, N., Weleff, J., & De Paoli, S. (2024). Inductive thematic analysis of healthcare qualitative interviews using open-source large language models: How does it compare to traditional methods? *Computer Methods and Programs in Biomedicine, 255*(10), 108356.

Mayer, C.-H., & Fouché, P. J. P. (2025). Psychobiography 4.0: About robots, AI and humanoids. *International Review of Psychiatry, 37*(5), 409−415.

Morgan, D. L. (2023). Exploring the use of artificial intelligence for qualitative data analysis: The case of ChatGPT. *International Journal of Qualitative Methods, 22,* 1−10. https://doi.org/10.1177/16094069231211248

Navigli, R., Conia, S., & Ross, B. (2023). Biases in large language models: Origins, inventory, and discussion. *Journal of Data and Information Quality, 15*(2), 1−21.

Nazir, A., & Wang, Z. (2023). A comprehensive survey of

ChatGPT: Advancements, applications, prospects, and challenges. *Meta-Radiology, 1*(2), 100022.

Perkins, M., & Roe, J. (2024). The use of generative AI in qualitative analysis: Inductive thematic analysis with ChatGPT. *Journal of Applied Learning and Teaching, 7*(1), 390−395.

Ponterotto, J. G. (2018). Advancing psychobiography: Reply to Young and Collins (2018). *American Psychologist, 73*(3), 288−289.

Rimban, E. L. (2023). Challenges and limitations of ChatGPT and other large language models. *International Journal of Arts and Humanities, 4*(1), 147−152.

Smirnov, E. (2025). Enhancing qualitative research in psychology with large language models: A methodological exploration and examples of simulations. *Qualitative Research in Psychology, 22*(2), 482−512.

Tschisgale, P., Wulff, P., & Kubsch, M. (2023). Integrating artificial intelligence-based methods into qualitative research in physics education research: A case for computational grounded theory. *Physical Review Physics Education Research, 19*(2), 020123.

Wyer, S., & Black, S. (2025). Algorithmic bias: Sexualized violence against women in GPT-3 models. *AI and Ethics, 5*(3), 3293−3310.

Xie, X., Shu, Y., He, C., Zhang, Y., & Ren, X. (2025). How individuals in collectivist cultures overcome crisis of self-development: A psychobiographical exploration of Li Qingzhao and her female self-awakening. *Journal of Constructivist Psychology*. https://doi.org/10.1080/10720537.2025.2561148

Zhang, H., Wu, C., Xie, J., Lyu, Y., Cai, J., & Carroll, J. M. (2023). Redefining qualitative analysis in the AI era: Utilizing ChatGPT for efficient thematic analysis. *arXiv*. https://doi.org/10.48550/arXiv.2309.10771

The application potential, challenges, and implications of artificial intelligence in psychobiography

SHU Yueyu, LI Chunjiang, REN Xiaoxiao, XIE Xia, ZHANG Yinxia, SONG Huan

(*School of Psychology, Northwest Normal University, Lanzhou* 730700, *China*)

Abstract: The rapid development of large language models (LLMs) in artificial intelligence presents new opportunities for psychobiography to address longstanding challenges, such as large data volumes, complex processing, and subjectivity. This paper systematically explores the potential, challenges, and implications of applying LLMs in psychobiography research. On the one hand, LLMs can improve research efficiency and analytical capacity. On the other hand, they also pose potential challenges, including data bias, interpretive deviation, and limitations in semantic understanding, as well as associated ethical risks. Facing these challenges, this paper proposes three possible solutions: first, constructing a localized large language model database to provide solutions to address cultural context and data bias issues; second, proposing ethical guidelines for human-machine collaboration in the field of psychobiography to guide the standardized development of psychobiography research in the era of artificial intelligence; third, proposing a theoretical concept of constructing a researcher-led human-machine collaborative research framework to point out the direction for the methodological path of psychobiography in the era of artificial intelligence.

Keywords: artificial intelligence, psychobiography, large language models, human-machine collaborative, DeepSeek

附录 A　本研究所使用的智能体及大语言模型

本文在分析时所使用的是基于ima-AI智能体的 DeepSeek V3.2 大语言模型。

之所以选择ima-AI智能体是因为该智能体能够由研究者根据自己的研究, 个性化的生成相关知识库。本文在使用过程中, 将《女性真实自我的寻求：以李清照的心理传记学分析为例》一文中所使用的分析资料单独建立了一个知识库, 使得后续的分析结果都是建立在这些资料的基础上, 从而最大限度的避免网络信息对分析结果的污染。

之所以选择 DeepSeek 作为分析的大语言模型, 是因为该文的研究对象为中国古代历史人物。而与其他大语言模型相比, DeepSeek 的中文理解与学术文本生成能力相对更适合本文的研究内容。

此外, 本文中所有运用DeepSeek进行的分析, 不同于以往实验研究中的实证分析模式, 而属于为本研究提供论据的“启发式示范”。

在评估大语言模型的结果与人类判断何者更优方面, 我们是基于其回答的真实性、逻辑性和创新性三个方面进行考量。例如, 在进行资料收集的过程中, 判断大语言模型收集的资料是否可靠时, 是通过对资料的真实性进行核对。在判断大语言模型推荐的理论是否恰当时, 通过将其推荐的理论与传主的生命故事相结合, 判断其解读的逻辑是否符合人类的正常思维方式。在识别理论盲点并提出创新点时, 是通过考察是否提出了原有资料中并不存在的新概念, 以判断该创新点是对已有资料归纳总结的“伪创新”还是真正具有原创性的创新。

附录 B　DeepSeek 提供的宋高宗相关资料

如以宋高宗为传主, 要查找有哪些可以研究的资料, 在 DeepSeek 中输入以下内容“你是一位专业的心理传记研究学者, 不仅精通心理学, 而且熟读中国历史。现在要写一部关于宋高宗赵构的心理传记, 给出可供参考的真实资料, 并给出每份资料的出处, 确保其属于学术专著。最后以表格的形式呈现出来” (附表 1)。经查证, 除《忧患与悲情：南宋初年的时代氛围与文人心态》和《皇权下的通权达变：宋高宗心理传记》两本书不存在外, 其余作品均是真实存在的。在两本不存在的书中, 前者找到了同一作者的相似书籍《宋代士风研究》, 后者则并未找到任何相似作品, 属于 AI 虚构的内容。

附表 1　DeepSeek 给出的宋高宗相关资料

类型	书名	作者/编者	出处
史料	《宋史》	脱脱等(元官修)	中华书局(1977)
	《建炎以来系年要录》	李心传(宋)	中华书局(2013 重印)
	《三朝北盟会编》	徐梦莘(宋)	上海古籍出版社(2019)
	《鄂国金佗稡编续编》	岳珂(宋)	中华书局(1989)
研究专著	《岳飞新传》	王曾瑜	中国书籍出版社(2016)
	《宋高宗传》	王曾瑜	中国书籍出版社(2016)
	《秦桧研究》	韩酉山	中华书局(2015)
	《南宋初期政治史研究》	[日]寺地遵	复旦大学出版社(2016)
	《忧患与悲情：南宋初年的时代氛围与文人心态》	郭学信	中国社会科学出版社(2021)
	《皇权下的通权达变：宋高宗心理传记》	何勇强	浙江大学出版社(2022)
	《宋光宗宋宁宗》	虞云国	吉林文史出版社(1997)

附录 C　DeepSeek 对文本内容倾向性的判断

运用 DeepSeek 分别对王曾瑜(2016)的《宋高宗传》和何忠礼(2021)的《宋高宗新论》进行模式识别和立场检测，指出作者对赵构的整体评价：

DeepSeek 对《宋高宗传》的分析：作者通过严谨的史料分析与批判性语言，构建了对赵构的负面主导评价，将其定位为“因私废公的妥协者”，同时承认其在特定历史条件下的有限作用。

DeepSeek 对《宋高宗新论》的分析：作者总体上持语境化的修正主义立场，试图在传统道德史观之外，建立以统治效能、文化延续性为核心的评价体系。这种立场既区别于全盘否定的"昏君论"，也不同于为尊者讳的官方史观，体现出当代史学对复杂历史人物的重新解读趋势。书中对心理因素、制度约束的强调，反映了现代史学方法对传统人物评价的革新。

DeepSeek对上述两书的分析结果与笔者的阅读感受和学界评论基本一致。如曹家齐(2022)认为《宋高宗新论》“该书在对宋高宗的人品、行事及历史功过是非的论述上，既有肯定，亦多否定，力图能以公允的立场和理性的价值观，展现宋高宗作为一代帝王及作为一个人的复杂性与多面性” (曹家齐, 2022)。

附录 D　DeepSeek 基于“真实自我”主题提供的理论备选

DeepSeek 基于“真实自我”主题的李清照心理传记学理论备选

一、针对“再嫁问题”的理论重构(聚焦真实自我冲突)

1. 温尼科特(D.W. Winnicott)的“真实自我/虚假自我”理论

虚假自我适应：宋代寡妇守节是理学推崇的“理想自我”，李清照再嫁可能是在病弱、流离状态下(《投启》称“牛蚁不分”)形成的虚假自我妥协——通过顺从社会规范(如寻求经济庇护)掩盖丧夫后的真实脆弱。

真实自我的反抗：婚后迅速揭发张汝舟“妄增举数”之罪并离婚，体现真实自我对虚假适应的摧毁——她宁可入狱(宋代妻告夫需徒二年)也不愿牺牲自主性，印证其“盐絮家风”中形成的核心自我(“诗情如夜鹊，三绕未能安”的独立人格)。

2. 罗杰斯(Carl Rogers)的“自我概念一致性”理论

再嫁与自我概念的冲突：李清照早期自我概念建立在“才女”“贤妻”身份(与赵明诚共研金石)，再嫁行为与这一概念严重不符，引发认知失调。离婚则是恢复自我一致性的行动(如《金石录后序》中“虽处忧患困穷而志不屈”的宣言)。

3. 埃里克森(Erik Erikson)的“认同危机”理论

晚年自我整合失败：50 岁后面临“自我整合 vs.绝望”阶段，再嫁是试图通过新关系重构认同(如扮演“受庇护者”)，但张汝舟的功利性(“觊觎文物”)迫使她直面自我真实需求——文物整理与文学创作才是其生成性(generativity)的核心(晚年致力《打马图经》编纂)。

附录 E　DeepSeek 对理论适配性的检验

运用检索增强生成技术检测李清照“为何再嫁”的悬疑性问题与上述三种理论的匹配度，DeepSeek 给出的建议是：

再嫁问题优先采用温尼科特虚假自我理论：能同时解释“妥协再嫁” (虚假自我适应)和“决然离婚” (真实自我回归)的矛盾行为。

对抗样本检测(Adversarial sample detection)技术，可以识别所选理论的失效场景，发现理论盲点，并提示可能的理论修正与发展方向。以下是 DeepSeek 给出的测试结果：

结果：理论通过测试，但需修正——温尼科特理论应加入文化语境(宋代寡妇的规范压力)。

但有趣的是，当让 DeepSeek 将其选出的“最佳”理论与研究者原有理论相比较后，却得出了以下结果：

原文理论更贴切，因其结合了社会文化(性别角色)与个体心理机制(补偿心理)，而第一组理论偏重普适性心理冲突，未充分纳入历史语境。

由此可知，大语言模型确实能够在整合文本资料、历史事件时间轴、社会文化数据等的基础上，提供与主题相契合的备选理论，帮助研究者缩小理论的选择范围。但大语言模型基于逻辑自洽性和数据匹配度的检验，固然能发现形式上的矛盾或证据缺口，但其“理论盲点检测”能力存在根本边界。