

# 业务建模驱动 TRIZ 注入的人-多智能体协作创新需求捕获框架<sup>\*</sup>



刘新华, 金 敏, 余梦蛟, 谢文涛

(湖南大学 计算机学院, 湖南 长沙 410082)

通信作者: 金敏, E-mail: [jinmin@hnu.edu.cn](mailto:jinmin@hnu.edu.cn)

**摘 要:** 当前软件市场呈现出产品同质化加重趋势, 功能性创新已成为决定软件竞争优势的关键因素. 这促使现代需求工程研究范式从被动的需求提取转向主动的创新需求捕获. 在提升需求创新性的实践中, 现有研究主要呈现两条路径: (1) 通过情景建模与引导方法改进工作坊流程, 激发人类团队协作创新; (2) 基于组合创新理论对既有需求进行解构重组, 快速生成新需求方案. 但两种方法均面临创新质量与参与成本难以有效平衡的核心矛盾. 生成式 AI 技术的突破性进展为应对这一挑战提供了新思路. 提出一种业务建模驱动下注入 TRIZ 理论的人-多智能体协作式创新需求捕获框架 BMHACT, 该框架以统一过程业务建模协作架构为蓝本, 设计提示词定义业务流程分析员、业务设计员等 5 个智能体职责. 多智能体团队通过“系统愿景收集-流程痛点识别-技术矛盾分析-TRIZ 创新原理匹配-需求方案生成”的协作流程生成创新需求, 并由领域专家和客户代表对需求进行创新性评估. 以工程机械领域某企业门户网站建设项目为例的实证研究表明: 相较基于需求重用的自动化方法和基于对抗样本的追溯式需求生成方法, BMHACT 迭代次数分别降低 50% 和 28.6%, 全过程耗时减少 66.7% 和 33.3%, 同时, 创新潜力指数 (clarity novelty usefulness, CNU) 分别提升 22.9% 和 10.7%, 且 CNU 单轮平均增益分别提高 2.16 倍和 2.14 倍. 证明了 BMHACT 在提升需求创新质量和降低协作成本上的优越性.

**关键词:** 创新需求; 大语言模型; 人-多智能体协作; 注入 TRIZ 理论的提示词工程; 多阶段质量门控机制; 创新潜力指数

中图法分类号: TP311

中文引用格式: 刘新华, 金敏, 余梦蛟, 谢文涛. 业务建模驱动TRIZ注入的人-多智能体协作创新需求捕获框架. 软件学报. <http://www.jos.org.cn/1000-9825/7562.htm>

英文引用格式: Liu XH, Jin M, Yu MJ, Xie WT. Business-modeling-driven Human-AI Multi-agent Collaborative Framework with TRIZ Infusion for Creative Requirements Capture. Ruan Jian Xue Bao/Journal of Software (in Chinese). <http://www.jos.org.cn/1000-9825/7562.htm>

## Business-modeling-driven Human-AI Multi-agent Collaborative Framework with TRIZ Infusion for Creative Requirements Capture

LIU Xin-Hua, JIN Min, YU Meng-Jiao, XIE Wen-Tao

(College of Computer Science and Electronic Engineering, Hunan University, Changsha 410082, China)

**Abstract:** The current software market is witnessing an intensified trend of product homogenization, where functional innovation has become a decisive factor in maintaining competitive advantage. This shift has transformed the paradigm of modern requirements engineering from passive requirements extraction to proactive creative requirements capture. Existing approaches to enhancing requirements creativity primarily follow two paths: (1) fostering collaborative innovation in workshops through scenario modeling and facilitation methods, and (2) rapidly generating novel solutions by deconstructing and recombining existing requirements based on combinatorial

<sup>\*</sup> 基金项目: 国家重点研发计划 (2023YFC3503404); 湖南省自然科学基金 (2025JJ50343)

收稿时间: 2025-04-21; 修改时间: 2025-06-10; 采用时间: 2025-09-29; jos 在线出版时间: 2026-01-21

innovation theory. However, both methods face a core challenge in balancing innovation quality with participation costs. The breakthrough advancements in generative AI technologies offer new opportunities to address this dilemma. This study proposes a business modeling-driven human-AI multi-agent collaborative framework with TRIZ infusion for creative requirements capture (BMHACT). The framework adopts the unified process business modeling collaborative architecture to design prompt-based definitions for five agent roles: business process analyst, business designer, and other relevant roles. The multi-agent team collaboratively generates creative requirements through a structured workflow: system vision collection→process pain point identification→technical contradiction analysis→TRIZ innovation principle matching→requirement solution generation. Domain experts and client representatives then evaluate the requirements for creativity. An empirical study on a portal system for a small-scale mechanical manufacturing enterprise demonstrates that, compared to the requirement reuse-based method and the adversarial-sample-based retrospective requirement generation method, BMHACT reduces iteration cycles by 50% and 28.6%, shortens total process duration by 66.7% and 33.3%, increases the clarity novelty usefulness (CNU) by 22.9% and 10.7%, and achieves a 2.16× and 2.14× higher per-round CNU improvement rate. These results validate BMHACT's superiority in enhancing requirements innovation quality while reducing collaboration costs.

**Key words:** creative requirement; large language model (LLM); human-AI multi-agent collaboration; TRIZ-infused prompt engineering; multi-stage quality gating mechanism; clarity novelty usefulness (CNU)

在软件工程领域,需求被定义为以自然语言描述的系统服务集合,其核心功能在于描述利益相关者的期望<sup>[1]</sup>.作为需求处理的核心环节,需求工程不仅是软件工程生命周期的核心阶段,更通过迭代式演进贯穿软件开发全流程.传统需求工程方法的核心任务在于捕捉利益相关者诉求,通过结构化方法(包括需求获取、分析与验证等环节)将模糊需求转化为可执行方案的技术契约,从而构建开发者与利益相关者的协作基础<sup>[2]</sup>.然而,这一方法在互联网经济背景下暴露显著局限性,主要体现在两方面:(1)仅依赖用户既有认知的显性需求提取,导致解决方案高度趋同,加剧产品同质化;(2)面对动态演化的用户需求(如新兴技术催生的场景、市场快速迭代的偏好),传统方法因缺乏主动创新机制而难以及时响应.这促使现代需求工程研究范式发生根本转变,即从被动的需求提取转向主动的创新需求捕获,通过系统性方法挖掘超越用户显性表达的潜在需求.

“创新需求”是指在软件系统开发过程中,为应对未来市场变革、技术发展趋势及用户潜在期望而提出的具有前瞻性、探索性和变革性的系统需求<sup>[3-5]</sup>.相较于传统需求,创新需求不仅解决当前问题,更致力于通过创新思维与前沿技术的应用,重塑用户体验并创造新的商业价值<sup>[6]</sup>.当前创新需求捕获的主流实践采用工作坊形式<sup>[7-14]</sup>,由资深分析师引导利益相关方通过头脑风暴和创造力支持工具生成需求.该方法在创新产出方面取得显著成效,但随着系统复杂度指数级增长和全球化协作的常态化,其在时间成本、地理限制及协作效率这3方面的局限性日益凸显.尤其在敏捷开发模式普及的背景下,需求响应周期缩短至数周,利益相关方和分析师团队因密集的会议安排与文档更新承受着严重认知负荷压力,制约了创新需求工程的规模化应用.为降低参与负担,一些学者提出自动化创新需求生成范式,但其实现路径仍面临显著挑战:(1)在语义推理路径中,基于领域本体<sup>[15]</sup>和主题建模<sup>[16,17]</sup>的方法依赖预定义本体的表征完备性,难以动态适配业务目标演化;(2)在过程建模路径中,隐马尔可夫模型<sup>[18]</sup>的过程建模因需求粒度过粗,导致生成方案实用价值有限;(3)在组合创新路径中,需求重用机制<sup>[19-21]</sup>虽遵循组合创造力<sup>[22]</sup>原则,但缺乏业务目标导向的筛选逻辑,易产生无效方案堆积.现有工作坊和自动化方法虽在特定场景可行,却均面临着创新质量与参与成本难以平衡的挑战.

生成式 AI 技术的突破性进展,特别是大语言模型 (large language model, LLM) 在深度推理与角色化需求生成能力的提升,为破解效率困境提供了新思路.靳东明等人<sup>[23]</sup>提出的虚拟需求分析团队框架 ChatModeler 在降低人力协作成本、提升需求建模效率方面展现出显著优势.该框架以人类需求提供者作为需求供给主体,通过提示词工程构建收集者、检测员、建模师等 LLM 代理角色,执行对话式需求捕获、模型元素提取及完整性验证等标准化处理流程,形成了人机协作的需求捕获范式.当前,该框架的技术路径主要聚焦于既定需求的规范化处理,在创新需求的生成机制构建及 LLM 跨领域知识重组、创造性推理潜能的挖掘利用方面,有待进一步探索.

本研究融合传统创新需求工程理论与 LLM 角色化任务执行的思路,提出业务建模驱动下注入 TRIZ 理论的人-多智能体协作式创新需求捕获框架 BMHACT.在统一过程 (rational unified process, RUP)<sup>[24]</sup>的业务建模阶段,通过识别系统的业务愿景与目标,深入分析现有业务流程中的瓶颈与改进机会,并在此基础上设计新的业务流程,由此生成具有业务语境支撑的创新性业务需求,实现从业务视角出发的需求创新.然而,在传统的 RUP 业务建模过

程中, 设计新业务流程及其映射为具体业务需求的过程, 往往依赖参与者的经验或直觉, 来应对诸如不同业务目标之间的冲突、资源约束与功能期望间的矛盾等问题. 在该过程中缺乏一个能够有效引导创新思维、结构化地解决矛盾并具备广泛适用性的系统方法论. 为弥补这一缺失并充分释放 LLM 的创新潜力, BMHACT 框架中引入了 TRIZ 理论<sup>[25]</sup>. 不同于依赖经验或直觉的头脑风暴, TRIZ 通过对全球跨领域专利的抽象凝练, 定义了 39 个通用工程参数和 40 条创新原理, 构建了一个普适的矛盾解决矩阵. 该理论赋能 LLM 将业务瓶颈/矛盾抽象为标准矛盾模型, 通过矛盾矩阵匹配创新原理, 驱动其发挥跨领域知识迁移与发散性创意生成能力, 输出创新性需求方案. 具体而言, BMHACT 框架包含三阶段流程. 首先, 基于 RUP 业务建模框架设计多角色提示模板, 构建需求收集员、业务流程分析员、业务设计员及系统分析员的多角色协同机制, 驱动 LLM 融合 TRIZ 创新原理<sup>[24]</sup>进行自主需求生成. 其次, 通过 GRUEN 评估模型<sup>[26]</sup>对生成需求实施语言质量过滤与逻辑一致性验证, 形成候选需求池. 最终, 由领域专家与客户代表使用五点李克特量表<sup>[20]</sup>从语义清晰性、功能新颖性、价值实用性这 3 个维度进行创新性评估. 未达创新潜力指数 (clarity novelty usefulness, CNU) 阈值的需求将通过人工反馈机制触发迭代优化, 达标需求则由需求建模员角色自动生成结构化需求文档. 此设计实现了“多智能体团队生成-人类验证反馈”的闭环协作范式, 能有效降低人员认知负荷, 提升创新需求的获取效率和质量.

为验证 BMHACT 对创新需求捕获效率与质量的提升效果, 本文以工程机械领域某企业门户网站建设项目为案例, 基于 DeepSeek-R1 模型构建智能体开展实证研究. 选取先进的自动化方法与 BMHACT 框架进行对比实验. 评估结果显示在生成相同数量创新需求的前提下, BMHACT 框架全过程耗时更短, 且需求创新性更高. 这证明了 BMHACT 可实现需求创新支持领域的“效率-质量”双维度优化.

本文的主要贡献包括 3 个方面.

(1) 提出了一种用于创新需求捕获的框架 BMHACT, 基于 LangChain 与提示词工程将 RUP 业务建模工作流程中的人员协作架构转化为多智能体协作机制, 实现系统愿景对话式获取、业务流程分析和业务设计等环节的自动化, 同时通过预设结构化文档提示词约束智能体输出, 确保需求规格的完整性与可追溯性.

(2) 设计了 TRIZ 与 LLM 深度协同的创新生成机制, 通过 LLM 从业务流程痛点中动态抽象高泛化性技术矛盾, 并基于 TRIZ 矛盾矩阵匹配相应创新原理<sup>[25]</sup>. 在此基础上, 利用提示词工程引导 LLM 运用内隐的跨领域知识, 将其与创新原理进行语义映射, 发散性生成兼具领域适配性与创新性的需求解决方案.

(3) 构建了“机器初筛-专家精评-闭环优化”的 3 层质量控制体系, 引入 GRUEN 语言质量评估模型进行需求的语言合规性与逻辑一致性检测, 形成候选需求池. 继而由领域专家与客户代表组成的评审委员会, 采用李克特量表进行创新性量化评估. 根据评审反馈与 LLM 因果追溯结果, 触发多智能体迭代优化流程.

本文第 1 节对相关工作进行介绍. 第 2 节说明本文所需的基础知识. 第 3 节介绍我们提出的框架和提示词设计细节. 第 4 节报告我们的研究问题和实验结果. 第 5 节对全文进行总结并说明未来的工作方向.

## 1 相关工作

### 1.1 工作坊方法

在创新需求捕获研究领域, 已有多种支持创新需求生成的工作坊方法被提出. Maiden 等人<sup>[8]</sup>提出的创意研讨会方法, 通过协调员引导下的头脑风暴激发参与者创造力. 随后, Maiden 等人<sup>[9]</sup>进一步提出了基于场景的需求工程方法 RESCUE 方法, 该方法采用场景驱动的结构化研讨框架捕捉创新需求. Jones 等人<sup>[10]</sup>则提出 APOSDLE 方法, 该方法采用为期两天的多阶段结构化流程 (含 4 个专项会议), 系统整合了 Boden 创造力理论<sup>[22]</sup>, 同时采用循环迭代、约束消除等技术同步激活探索性、组合性与变革性这 3 种创造力. 此外, Karlsen 等人<sup>[11]</sup>设计了集成软件工具以促进信息采集中的创新思维. Sakhnini 等人<sup>[12]</sup>则引入基本语用模型 (elementary pragmatic model, EPM) 技术, 扩展了创意搜索空间. Ferrari 等人<sup>[13]</sup>提出的歧义驱动审查法旨在通过将审查前移至需求获取阶段以揭示客户隐性知识, 进而提升需求的新颖性与完整性. Giunta 等人<sup>[14]</sup>提出创造力触发机制, 基于发散-聚合思维模型<sup>[27]</sup>, 通过预设创新品质卡片触发利益相关者的创造力. 然而, 这些方法均依赖大量人工投入, 未能实质性降低需求提供方的认知负荷, 且实施效果持续受制于参与主体的创造潜能和协调者的经验水平.



## 1.2 自动化方法

在自动化创新需求生成方法研究领域,学者们已通过多维度探索形成系列研究成果。Farfeleder 等人<sup>[15]</sup>构建了基于领域本体的语义引导系统,通过形式化定义概念、关系及公理体系实现创新需求自动化生成。Bhowmik 等人<sup>[16]</sup>则提出组合式需求生成框架,基于组合创造力理论<sup>[22]</sup>,运用主题建模技术提取利益相关者讨论的共性主题,结合词性转换机制生成跨领域概念组合,从而自动发掘潜在创意触发点。Do 等人<sup>[18]</sup>在后续研究中引入隐马尔可夫模型(hidden Markov model, HMM),通过挖掘需求文本中的潜在创造性特征实现新需求推导。随后,Do 等人<sup>[20]</sup>进一步突破传统范式,设计出基于在线软件需求数据的需求重用框架,借助跨领域组合策略提升创新需求的生成效率。上述基于组合创造力理论的自动化方法,在提升创新需求捕获效率方面取得显著进展,但其在需求创新质量方面的局限性逐渐显现。生成过程存在显著随机性,导致输出需求在语义连贯性与逻辑完整性方面存在缺陷,仍需依赖领域专家进行人工验证与修正。同时,由于技术多局限于组合性创造力<sup>[22]</sup>路径,创新机制受限于既有需求知识库的边界约束,生成结果多呈现已有元素的线性组合特征,在突破性创新方面尚未形成有效解决方案。这种路径依赖现象使得生成需求难以超越既定知识体系,制约了具有颠覆性创新价值的软件需求的产生。

而 Gudapartihi 等人<sup>[21]</sup>将对抗生成的研究思路引入本领域,提出基于对抗样本追溯的需求生成方法。该方法通过对原始需求样本进行词级扰动构造对抗样本,凭借 GRU 分类器对对抗样本与原始样本组成的平衡数据集进行迭代分类,依据分类器误判结果及排名机制动态调整扰动位置,实现“扰动-分类”的协同进化,在满足词性标注约束前提下提升需求创新性。

## 1.3 需求创新性评估方法

在需求工程研究领域中,创新性概念的操作化定义已形成以“新颖且有用”为核心特征的理论共识<sup>[19,27-30]</sup>。针对该概念的量化评估,学界发展出多维度、多方法的测量体系:在量表设计层面,Maiden 等人<sup>[8,9]</sup>在航空管理系统功能需求评估中首创双专家独立评审机制,采用 3 点李克特量表分别测量“新颖性”与“适用性”;在分类方法层面,Karlsen 等人<sup>[11]</sup>在安全访问系统研究中则应用二元分类法(创意/非创意)实现快速评估。值得注意的是,Zachos 等人<sup>[31]</sup>在 Web 服务需求研究中率先引入文本相似度算法,为创造力的客观量化提供了技术路径。在评估指标体系的演进过程中,Bhowmik 等人<sup>[16,19]</sup>采用单维度李克特量表对“新颖性与实用性”进行综合测量,而 Murukannaiah 等人<sup>[29]</sup>通过大规模实证研究( $N=300$ )构建了包含清晰性、新颖性与实用性的三维评估框架,该框架亦被 Do 等人<sup>[20]</sup>的研究中应用并优化。就评估主体而言,Burnay 等人<sup>[30]</sup>采用非专家评估范式,基于 8 名学生的反馈验证创造力触发器的有效性。相较之下,Giunta 等人<sup>[14]</sup>设计了三阶段混合方法研究,整合项目赞助商(1 人)、硕士生(19 人)与业务分析师(10 人)这 3 类利益相关者,通过创意数量统计(完整性)、新颖性标记评分(质量)与半结构化访谈(采纳意愿)实现多维数据三角验证。之后,Bhowmik 等人<sup>[32]</sup>在需求工程创造力评估框架研究中取得新进展,该团队创造性地构建了包含 4 个维度 27 项指标的 CRDS 量表(创造性需求诊断量表),并通过指标-需求活动映射机制首次实现创造力评估与需求工程全生命周期的深度关联。

本研究严格遵循需求导向的创造力评估范式,考虑到评估方法的轻量化与易实施性,在方法论层面继承并发展了 Do 等人<sup>[20]</sup>的研究框架,着重考察需求规格说明的清晰性、新颖性与实用性这 3 个核心属性,并在此基础上提出综合量化需求创新性的指标 CNU。需要特别指出的是,Maiden 等人<sup>[17]</sup>通过结构化工作坊( $N=10$ )建立的“机会新颖性”评估体系,包含独特性、领域突破性与行业异质性这 3 个子维度,为本研究的指标设计提供了重要的理论参照。

## 2 基础知识

### 2.1 RUP 中的业务建模工作流程

在软件工程方法论的发展过程中,RUP 首次系统化提出了“业务建模先行”原则,将业务建模确立为需求分析的前置流程,强调在软件开发的初期明确业务目标和流程,确保需求分析始终从业务目标出发,而非单纯依赖传统的用户调研和需求收集活动,这种传统方法往往造成开发过程中的需求误差与偏离<sup>[33,34]</sup>。引入业务建模流程的核心目标是通过结构化方法准确识别和分析现有业务流程,定义优化后的目标业务模型,并挖掘潜在的业务创新机

会,从而为后续的需求分析奠定坚实基础.这一业务驱动的需求分析方法体现了 Ivar Jacobson、Grady Booch 和 James Rumbaugh 这 3 位面向对象方法学大师对“业务-需求”共生关系的深刻洞察,强调了需求分析与业务目标和流程的紧密关联.这一理念深刻影响了现代敏捷开发中的用户故事映射,并为领域驱动设计 (DDD) 提供了重要的理论支持,特别是在如何将业务目标与软件设计有机结合方面.

具体而言,业务建模工作流程从项目初始阶段启动,通过访谈、文档分析或跨部门工作坊识别业务参与者及其目标,构建业务用例模型,以描述业务系统与内外部角色(如客户、供应商、部门)之间的交互逻辑<sup>[24]</sup>.这一阶段通过解构现有业务流程,能够直观暴露业务模式中的低效环节,从而为创新需求的提出奠定基础.随后,在目标业务流程的设计中,业务建模通过引入业务规则清单和创新场景假设,鼓励利益相关者突破既有框架,探索技术赋能的业务模式重构<sup>[35]</sup>.业务建模的成果(如业务用例文档、规则清单)通过迭代评审与持续协作,确保技术团队能够精准识别并实现具有突破性的需求,从而将创新理念转化为可落地的系统功能,最终推动业务价值的跃升.

## 2.2 TRIZ 创新理论

TRIZ 创新理论是由 Altshuller 等人<sup>[36]</sup>于 20 世纪 40 年代提出的系统性创新方法论,其通过分析海量专利与工程案例,提炼出技术系统演化规律及普适性矛盾解决原则,以突破传统试错法的局限<sup>[25]</sup>.TRIZ 理论体系包含 5 大核心工具:技术矛盾与物理矛盾分析、40 条创新原理、技术系统进化法则、物场模型以及最终理想解<sup>[37]</sup>.其中,技术矛盾指系统改进某一参数时导致另一参数恶化的对立关系(如强度提升导致重量增加),而物理矛盾则描述同一参数需同时满足相反需求(如温度需高又需低).通过矛盾矩阵与 40 条创新原理(如分割、抽取、预先作用等),TRIZ 为创新需求捕获过程提供了结构化路径以消除矛盾并生成突破性方案.该理论作为被广泛验证的创新工具,已成功应用于汽车<sup>[38]</sup>、公共服务<sup>[39]</sup>、纺织设计<sup>[40]</sup>等多个行业领域的创新解决方案开发.相较于传统头脑风暴,TRIZ 理论通过逻辑化、可复用的创新模式,显著提升创新需求捕获过程的系统性与可落地性.

## 2.3 大语言模型与提示词工程

LLM 是基于 Transformer 架构的深度学习模型,通过海量文本数据的预训练与微调,具备理解自然语言、生成连贯文本及执行复杂推理任务的“涌现能力”<sup>[41]</sup>.其核心特征在于对语义关联的隐式建模,能够从输入文本(即提示词)中捕捉上下文逻辑,并生成符合语法与领域知识的内容<sup>[42]</sup>.提示词工程作为引导大语言模型输出的关键技术,通过设计结构化提示(如角色设定、示例参考、约束条件),将开放式需求转化为模型可解析的指令,从而激发其创新潜力<sup>[43]</sup>.

相较于传统需求挖掘方法,大语言模型的零样本学习与上下文学习特性使其能够突破人类经验局限,快速生成多样化解决方案.而提示词工程通过设计输入模板(如类比推理、反事实假设),可系统性探索需求空间中的潜在创新点.两者的协同作用为需求生成提供了可扩展、高迭代的智能化工具链.

# 3 业务建模驱动下注入 TRIZ 理论的人-多智能体协作式创新需求捕获框架

## 3.1 BMHACT 框架与需求创新性产生机制概述

本文在借鉴传统需求捕获方法的多主体协作范式的基础上,融合 RUP 业务建模工作流程<sup>[24]</sup>与生成式人工智能技术,提出 BMHACT 框架.如图 1 所示,该框架通过智能体角色重构与认知增强机制实现范式突破,在保留传统方法中客户-专家双核评估架构的基础上,构建由 5 类智能体(需求收集员、业务流程分析员、业务设计员、系统分析员及需求建模员)组成的协作网络.

框架运作机制呈现螺旋式迭代特征:初始阶段由需求收集员智能体通过对话式交互捕获客户原始诉求;业务流程分析员基于系统目标锚定核心业务流程,联网执行行业流程数据分析与量化评估,定位流程痛点;业务设计员生成基于创新原理的需求方案;系统分析员将其转化为用例模型并进行质量验证;最终由需求建模员完成标准化文档输出.领域专家团队从清晰性、新颖性和实用性这 3 个维度实施终审,未通过评估的需求将触发反馈驱动的迭代流程,生成优化建议进入下一迭代周期.

BMHACT 框架生成需求的新颖性来源于 3 个方面:(1)摒弃传统场景构建与头脑风暴方法,在 RUP 业务建模指导下,基于系统目标锚定核心业务流程,通过行业流程数据分析与量化评估,定位高价值可改进节点.(2)基于

TRIZ 矛盾分析理论<sup>[25]</sup>, 首先对改进节点的技术矛盾进行识别与分类, 进而通过矛盾矩阵推导适配的创新原理, 并将该原理集作为业务设计的核心引导条件. (3) LLM 自身所具备的跨领域知识迁移、发散性创意生成能力.

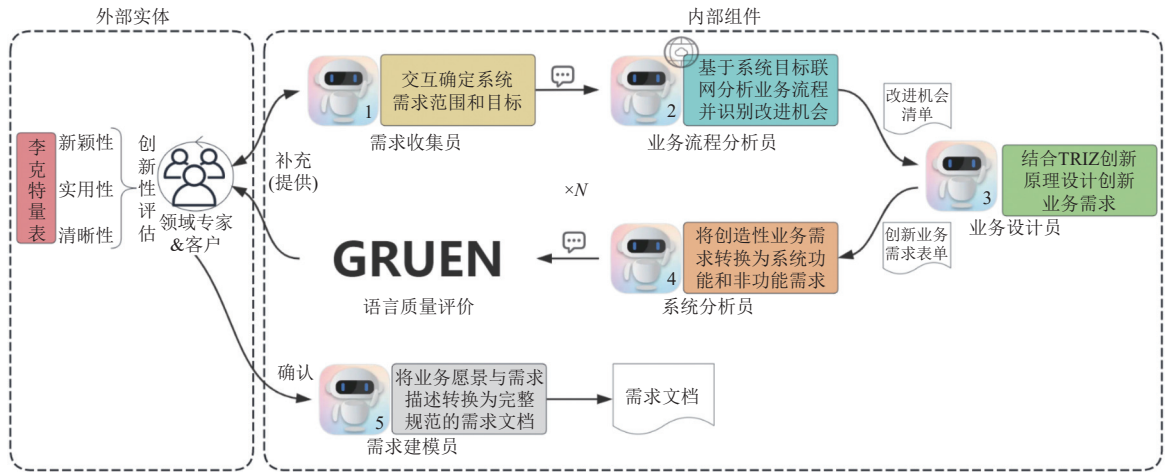


图 1 BMHACT 框架

3.2 BMHACT 框架的人-多智能体协作流程

本文提出的 BMHACT 框架通过 LLM 代理的多角色链式协作实现创新需求迭代优化, 其协作流程如图 2 所示, 可分为 6 个连续阶段.

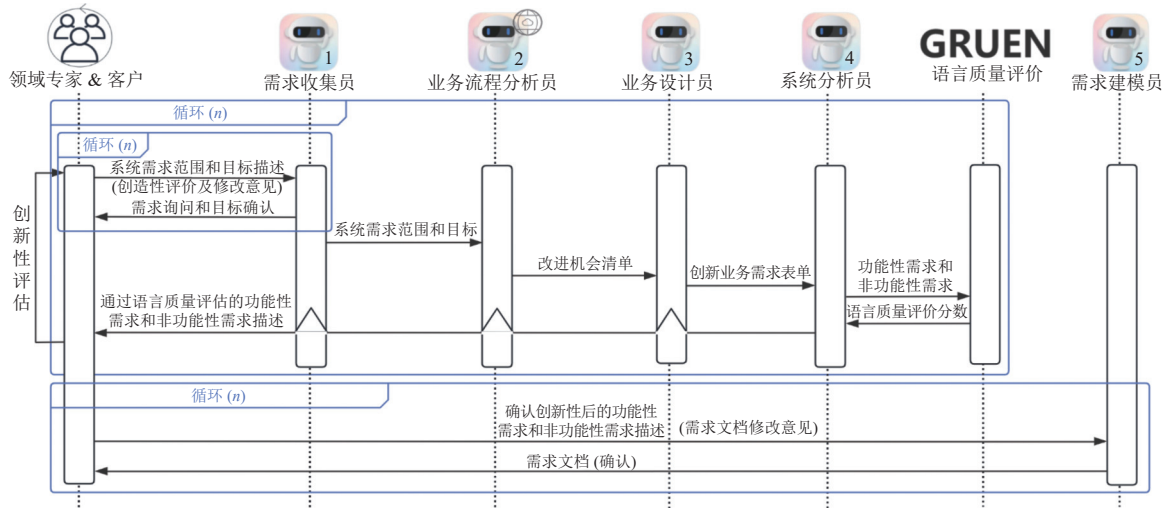


图 2 人-多智能体协作图

(1) 需求对话建模阶段

LLM 扮演的需求收集员通过多轮对话交互捕捉客户原始需求, 运用上下文理解技术动态跟踪对话状态, 将非结构化对话内容转化为包含系统边界定义、核心目标陈述和约束条件说明的结构化需求范围说明书.

(2) 市场分析与改进点识别阶段

此阶段由 LLM 扮演的业务流程分析员主导. 如图 3 所示, 业务流程分析员基于前序需求范围说明书启动联网检索模块, 获取同行业该业务流程数据并构建多维度对比分析矩阵, 通过“流程时间-负面评价”量化评估识别出具有改进价值的业务流程关键节点, 最终生成包含优先级排序的改进机会清单, 为后续创新提供定向输入.

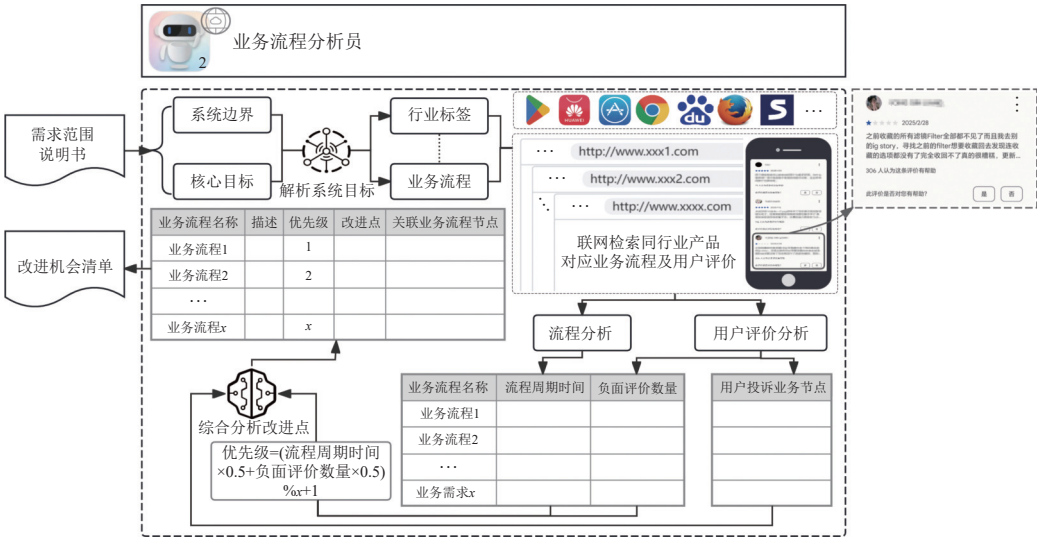


图3 业务流程分析员任务细化图

(3) 创新需求生成阶段

如图4所示,由LLM代理的业务设计员采用TRIZ矛盾分析框架解析改进机会清单中的核心冲突,结合创新原理发散性思考生成候选创新方案,通过预定义的“可行性-价值评估”策略对方案进行筛选。此阶段融合了创新启发式规则库(含40条创新原理及其最佳实践)和LLM的创新生成能力,在可控范围内探索需求创新空间,最终生成创新业务需求表单。完整的TRIZ矛盾矩阵及创新启发式规则库详情见<https://gitee.com/lxh8888/bmhact>。

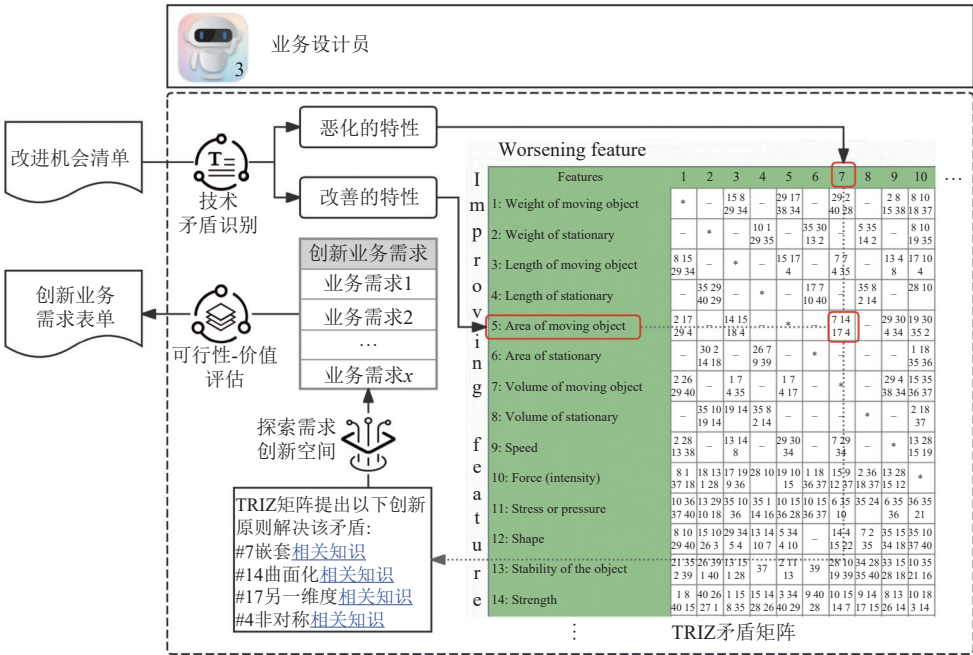


图4 业务设计员任务细化图

(4) 需求工程化验证阶段

由LLM扮演的系统分析员对创新业务需求进行结构化解析,识别其中的功能需求与非功能需求,并调用



GRUEN 模型<sup>[26]</sup>检测需求描述的语言质量, 最终生成包含用例和非功能需求描述的规格化需求描述文档.

(5) 多维度需求评审阶段

由领域专家和客户组成的需求评估委员会, 采用三维评估矩阵 (清晰性、新颖性、实用性) 的五点李克特量表对需求方案进行加权评分. 评估过程融入对抗性验证机制, 通过“人类专家质疑-大模型反演响应”迭代提升评估深度, 最终生成包含改进建议的评审报告.

(6) 闭环迭代优化阶段

根据评审报告结果动态选择路径. 通过评审的需求项进入文档生成流程, 由需求建模员按照提示模板生成可追踪的最终需求文档; 未通过项则启动反馈驱动的迭代优化, 直至满足终止条件 ( $CNU \geq 4.5$  或最大迭代次数阈值 10).

本研究在第 3.3 节针对流程涉及的中间件文档进行了系统化定义, 其技术路径为基于提示词工程技术构建标准化模板, 实现文档自动生成机制. 该方法通过预定义语言规则与结构约束, 确保中间件文档生成过程的规范性与一致性, 从而形成可复用的方法论保障体系.

3.3 多智能体协作团队构建

3.3.1 团队任务定义

基于第 3.2 节设计的角色协作流程, 本研究受启发于靳东明等人<sup>[23]</sup>提出的团队协作提示框架, 构建了面向创新需求捕获任务的团队提示词模板. 如表 1 所示, 该模板由 4 个核心要素构成.

表 1 创新需求捕获团队角色提示词模板

参数	具体说明
Team	现有一个创新需求捕获团队, 包括1个需求收集员、1个业务流程分析员、1个业务设计员、1个系统分析员和1个需求建模员, 不同人员之间分工不同, 相互合作
Role	我希望你能在这个需求团队中扮演一个<Role>, 负责<Duty>. 你会收到来自<Team> 的<Message>
Action	如果你收到<Team>的<Message>, 请做<Action>
Message	来自<Team>的<Message>的内容是<Content>

注: Prompt=Team+Role+Action+Message. Role表示大模型扮演的角色, Duty表示角色的职责, Team表示发生交互的其他角色, Message表示交互角色会发送什么信息, Action 表示角色执行的动作, Content表示来自交互角色的反馈内容. 这些都是提示词模板的参数

- (1) 团队架构说明 (Team): 描述多智能体协作体系, 明确各角色间的协同约束与交互范式.
- (2) 角色职能规范 (Role): 界定 LLM 需承担的具体角色及其知识边界, 并建立角色-职责的映射关系.
- (3) 行为操作指令 (Action): 严格定义 LLM 在接收跨角色信息后应触发的响应策略与执行路径.
- (4) 交互信息流 (Message): 采用对话历史顺序存储机制, 实时维护多角色间的对话上下文与知识共享状态.

这种分层递进的设计范式不仅实现了协作语义的精确表征, 更通过模块化架构增强了提示词工程的可解释性与可扩展性.

3.3.2 智能体领域知识增强

为提升智能体在需求捕获过程中的专业性和规范性, 本文构建了软件工程领域知识库. 该知识库基于 Rational 统一过程框架, 聚焦其业务建模与需求捕获中的角色定义、任务流程及最佳实践. 具体而言, 我们选取两本 RUP 著作<sup>[24,44]</sup>中定义的需求捕获相关角色、任务流程、建模规范及最佳实践作为知识源, 通过嵌入模型进行文本分割与向量化处理, 并采用 FAISS 构建向量索引, 实现高效的相似性检索. 所有智能体决策时自动执行实时检索, 筛选 Top-5 最相关的 RUP 核心概念要素 (如 Use Case 扩展点、非功能性需求约束等), 作为领域知识增强的决策依据.

此外, 本文还依据各智能体所承担的角色职责进行了针对性功能增强. 为业务流程分析员配备了联网搜索能力以实时获取行业信息; 为业务设计员集成了 TRIZ 矛盾矩阵及创新原理知识库, 以支持系统性创新; 同时, 为所有智能体均设计了标准化的输出中间件提示词模板, 以规范其输出内容与格式.

3.3.3 构建多智能体间上下文共享的协同进化机制

本文基于 LangChain 框架从“纵向-横向”两维度实现多智能体间的协同进化机制.

在纵向维度上, 依托 RunnableLambda 模块构建链式多智能体协作管道, 形成具有时序依赖的协作流程. 该流



程强制执行“需求收集员→业务流程分析员→业务设计员→系统分析员”的顺序执行逻辑. 下游智能体自动继承上游输出的状态作为输入基线, 在上游状态未就绪时, 下游智能体将自动阻塞, 从而确保整个创新路径在演化过程中始终处于原始需求语义空间之内, 避免信息偏移和理解失真.

在横向维度上, 基于 ChatMessageHistory 组件构建统一的全局状态容器, 并通过角色标识和时序标记机制实现 5 阶状态分层: 原始需求层、改进机会清单层、创新方案层、用例文档层以及需求文档层. 每个状态层对应特定智能体的执行环节, 系统将各智能体包含角色标识和时间戳的推理过程及其生成的中间产物进行分层持久化存储至全局状态容器中. 在状态共享方面, 通过 MessagesPlaceholder 模块动态加载全局完整对话历史至各智能体上下文, 构建全周期可追溯的知识传承链. 例如, 当人类评审专家输入需求修改指令开启新一轮迭代时, 智能体可即时调取历史状态, 结合反馈意见与上游输出状态进行增量式修正, 形成闭环质量控制回路.

3.3.4 智能体提示词工程

本文提出的多智能体创新需求捕获团队 (需求收集员、业务流程分析员、业务设计员、系统分析员、需求建模员) 在创新需求工程中的协同机制与阶段分工, 已如本文第 3.2 节所述. 从提示词工程的技术实现层面, 各角色的智能体构建实质上是对第 3.3.1 节所述团队协作提示模板的实例化演绎. 以业务设计员角色为例, 其提示模板定义包含 Team、Role、Action 及 Message 这 4 个核心要素. 如表 2 所示, 在 Role 定义中, 我们通过提示词构建了 TRIZ 矛盾分析驱动的创新路径, 明确要求智能体基于 TRIZ 理论解构业务流程可改进点; 在 Action 流程中, 则系统地植入了 TRIZ 与 LLM 协同创新机制, 通过“技术矛盾分析-创新原理匹配-需求方案生成”的递进式 workflow, 将 TRIZ 方法论深度融入业务设计环节. 其他角色的具体实例化提示模板详见: <https://gitee.com/lxh8888/bmhact>.

表 2 TRIZ 创新理论引导的业务设计员智能体构建提示词

参数	取值
Team	现有一个创新需求捕获团队, 包括1个需求收集员、1个业务流程分析员、1个业务设计员、1个系统分析员和1个需求建模员, 不同人员之间分工不同, 相互合作
Role	我希望你能在这个需求团队中扮演一个<业务设计员>, 负责<运用TRIZ矛盾分析框架解析改进机会清单中的核心冲突, 结合40条创新原理发散生成候选方案, 通过可行性-价值评估策略筛选创新方案, 最终生成创新业务需求表单>. 你会收到来自<业务流程分析员>的<改进机会清单>
Action	如果你收到<改进机会>, 请按顺序执行: <1. 使用TRIZ矛盾分析框架拆解清单中的核心冲突, 识别技术矛盾与物理矛盾> <2. 根据矛盾类型匹配创新原理库 (40条原理及对应最佳实践), 生成3-5个候选创新方案> <3. 应用预定义的可行性-价值评估策略 (技术可行性/ROI/实施成本三维评分)筛选方案> <4. 整合优选方案按预定义中间件提示词形成创新业务需求表单, 包含创新业务需求描述、技术解决方案和关联业务流程节点>
Message	来自<业务流程分析员>的<改进机会清单>的内容是<Content>

注: Prompt=Team+Role+Action+Message; Content表示来自业务流程分析员的改进机会清单

3.3.5 中间件文档提示词工程

本文提出的 BMHACT 框架在人-多智能体协作机制中涵盖的核心中间件文档包含需求范围说明书、改进机会清单、创新业务需求表单、需求描述文档及最终需求文档等多维产出物. 相关中间件文档依据提示词工程技术进行了系统化定义, 通过构建标准化提示词模板, 实现符合预设规范的文档自动生成机制. 该技术路径通过预定义语言规则与结构约束, 能有效确保文档生成过程的规范性与输出结果的一致性. 其中, 最终需求文档的提示模板如下所示, 采用分层结构化设计, 包含 7 个核心部分.

- (1) 概述部分: 定义系统名称、开发目标及背景动机.
- (2) 角色定义部分: 通过角色-职责-目标三元组明确利益相关者图谱.
- (3) 创新功能部分: 以名称-描述-价值-实现四维结构独立定义创新功能特征.
- (4) 功能需求部分: 基于用例模型 (Use Case) 描述原子化操作流程, 含主或替代路径及前后置条件.
- (5) 质量属性部分: 按 FURPS+框架 (功能性、可靠性、性能、可支持性) 定义非功能需求.
- (6) 约束模块: 声明技术兼容性、法律合规性等边界条件.
- (7) 附录模块: 包含术语标准化定义与文献引用, 支撑需求可追溯性.

---

请按照以下模板生成创新需求文档, 注意只输出文档内容, 不要包含其他额外信息:

## ## 1. 概述

### ### 1.1 系统名称

<简要描述系统的名称, 例如“智能库存管理系统”>.

### ### 1.2 系统开发目标

<描述系统需要实现的主要目标, 例如“提高库存管理效率, 减少存货不足和过量的风险”>.

### ### 1.3 背景与动机

<解释开发该系统的背景及原因, 例如“当前库存管理依赖手动操作, 容易导致库存不准确”>.

## ## 2. 系统角色

列出系统中涉及的主要角色, 包括每个角色的职责和目标:

\*\*\*角色名称:<例如“仓库管理员”>.\*

- \*\*职责: \*\* <描述该角色在系统中的具体职责, 例如“更新库存记录, 处理入库和出库请求”>.

- \*\*目标: \*\* <描述该角色通过系统实现的目标, 例如“确保库存数据实时更新”>.

## ## 3. 创新功能点

列出系统中引入的具有创新的新功能, 明确其价值和实现方式:

\*\*\*功能点名称:<例如“库存智能预测”>.\*

- \*\*描述: \*\* <简要描述功能点, 例如“基于历史数据和销售趋势预测未来库存需求”>.

- \*\*价值: \*\* <说明功能点对系统的改进价值, 例如“减少存货不足的情况, 提高库存周转率”>.

- \*\*实现方式: \*\* <简要说明功能的实现途径, 例如“通过机器学习模型进行需求预测”>.

## ## 4. 功能性需求

按照用例格式详细描述功能性需求:

### 用例编号: <唯一标识编号, 例如“UC001”>.

#### 用例名称: <简要描述用例, 例如“产品入库”>.

##### 描述: <描述系统为用户实现的功能, 例如“系统允许仓库管理员录入新到货物的相关信息”>.

##### 主要操作流程:

1. 用户 <描述用户的具体操作, 例如“选择‘入库’功能, 输入货物信息”>.

2. 系统 <描述系统的响应, 例如“验证货物信息并更新库存数据”>.

3. 用户 <描述后续操作, 例如“确认入库完成”>.

4. 系统 <描述最终动作, 例如“生成入库记录”>.

##### 替代流程 (如有):

1. 用户 <描述可能的替代操作, 例如“扫描条码直接录入货物信息”>.

2. 系统 <描述系统对此的响应, 例如“从条码中提取信息并更新库存数据”>.

##### 前置条件: <描述用例执行的必要条件, 例如“用户必须是经过授权的仓库管理员”>.

##### 后置条件: <描述用例完成后的状态, 例如“库存记录更新成功”>.

## ## 5. 非功能性需求

\*\*\*性能要求: \*\* <描述性能需求, 例如“支持每秒 1000 次库存查询请求”>.

\*\*\*可靠性要求: \*\* <描述可靠性需求, 例如“系统在断电恢复后数据无丢失”>.

\*\*\*安全性要求: \*\* <描述安全性需求, 例如“所有用户登录需要多因素认证”>.

\*\*\*可维护性要求: \*\* <描述可维护性需求, 例如“代码模块化设计, 支持快速升级”>.

## ## 6. 系统约束条件

列出系统设计和实现需要满足的约束:

---

\*\*\*技术限制: \*\* <描述技术约束, 例如“系统必须兼容现有的 ERP 系统”>.

\*\*\*法律与合规性: \*\* <描述合规性要求, 例如“遵守《数据保护法》确保用户隐私”>.

## 7. 附录

\*\*\*术语表: \*\* <列出文档中提到的关键术语及其定义, 例如“库存周转率: 指库存在一定时间内的循环次数”>.

\*\*\*参考文献: \*\* <列出文档引用的相关资料, 例如“Smith, J. (2023). Inventory Optimization Techniques”>.

注: Prompt=Explain+Document

文档架构兼容 ISO 29148 标准, 在需求陈述层 (第 1、2 部分) 与需求规格层 (第 3–6 部分) 间设置创新功能隔离带, 通过 Markdown 语义标记实现人机可读性统一. 功能性需求采用扩展用例模板 (含操作流程、异常处理及状态约束), 非功能性需求覆盖性能、安全等质量属性, 约束条件明确定义系统实现边界. 其他中间件文档的提示模板详见: <https://gitee.com/lxh888/bmhact>.

3.4 创新性评估量表

本文严格遵循需求导向的创新性评估范式, 在方法论层面继承并发展了 Do 等人<sup>[20]</sup>的研究框架, 着重考察需求的清晰性、新颖性与实用性这 3 个核心属性. 本文设计了一份三维度评估问卷, 如图 5 所示, 该问卷旨在对创新需求进行多维度特征解构, 其核心评估维度包括: 语义清晰性 (clarity)、需求新颖性 (novelty) 与价值实用性 (usefulness). 各维度操作化定义如下.

创新性评估表

评分说明: 1=完全不同意, 2=不同意, 3=中立, 4=同意, 5=完全同意

属性	问题	1	2	3	4	5
清晰性 (clarity)	该需求易于理解无歧义	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	该需求逻辑自洽且无语法错误	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
新颖性 (novelty)	该需求体现了独创性和意想不到的创新	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	该需求不同于现有的常见解决方案	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
实用性 (usefulness)	该需求能为用户提供实际价值	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	该需求在技术上可实现	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

图 5 创新性评估五点李克特量表

(1) 清晰性 (clarity): 该维度衡量需求描述的确定性水平与细节完备程度.

(2) 新颖性 (novelty): 该维度着重评估需求的原创突破性与预期超越性, 即需求方案是否突破常规认知边界并产生意外性价值.

(3) 实用性 (usefulness): 回归需求工程本质, 该维度衡量需求在系统生态中的实用价值, 着重考察其与系统目标的契合度及利益相关者价值实现潜力.

在评估方法学层面, 本研究采用专家-客户协同验证机制. 由领域专家与客户代表组成评估委员会, 基于五点李克特量表 (1=完全不同意; 5=完全同意) 进行双盲评分. 量表设计严格遵循心理测量学规范, 各题项均通过 Cronbach’s  $\alpha$  系数验证其内部一致性 ( $\alpha>0.85$ ), 并通过探索性因子分析 (EFA) 显示因子载荷大于 0.7 确保了维度间区分效度. 值得关注的是, 本研究在 Do 等人<sup>[20]</sup>的需求创新性评估框架基础上, 提出创新潜力指数 (CNU), 通过三维度得分加权计算 ( $CNU=0.3\times clarity+0.4\times novelty+0.3\times usefulness$ ), 为需求的创新性评估提供综合性量化指标.

4 工程项目开发案例实验

本文以工程机械领域某企业门户网站实际建设项目为研究对象,验证所提框架的实用价值.项目核心功能包括以下模块:面向研发团队的新产品信息管理系统;针对销售部门的新媒体资讯发布平台等基础模块.值得注意的是,为构建差异化竞争优势,客户明确提出需突破行业常规,开发竞品网站尚未实现的创新功能模块.该项目具有典型行业特征与创新属性:(1)项目真实存在创新需求挖掘的挑战,能够有效检验方法论的实践价值;(2)企业实体与业务场景的强关联性,为业务建模方法提供了理想的验证环境.

在实证研究过程中,特别设置对比实验,将本文提出的 BMHACT 框架与基于需求重用的自动化方法和基于对抗样本的追溯式需求生成方法进行对比分析.这种横向比较不仅能验证本方法在创新需求捕获方面相较先进自动化方法的表现,也将揭示不同方法论在实际应用场景中的适应性差异.

为了评估 BMHACT 框架的有效性,在实验开始前,我们提出如下问题.

- Q1: 使用 BMHACT 框架捕获创新需求所需的时间是否明显少于基线方法?
- Q2: 使用 BMHACT 框架生成需求创新质量 (即创新潜力指数) 是否明显高于基线方法?

4.1 实验设置

(1) 评估团队构成. 本研究组建跨学科评估团队,由 2 名领域专家 (来自本研究机构的技术团队) 与 4 名客户代表 (含 2 名企业代表: 研发部主管与销售部主管, 以及 2 名工程机械设备意向买家代表) 构成. 所有参与人员均签署保密协议并进行匿名化处理. 实验采用双盲评审机制,要求专家与客户代表分别从清晰性、新颖性、实用性这 3 个维度独立评分,评分工具为第 3.4 节设计的五点李克特量表.

在实验启动前,面向评估团队系统性阐述评估指标体系与案例约束条件,包括评分量表的解读,评估维度权重的设定原理,案例特异性参数 (技术场景、业务边界). 通过此流程确保参评人员对评估标准达成共识.

(2) 对比实验设计. 为验证 BMHACT 框架的有效性,选择基于需求重用的自动化方法<sup>[20]</sup>和基于对抗样本的追溯式需求生成方法<sup>[21]</sup>作为基线,比较其与 BMHACT 在创新需求生成效率及质量方面的表现.

(3) LLM 选型、部署与应用策略. 采用上海人工智能实验室发布的司南评测体系 (OpenCompass 2.0) 进行模型选型. 该体系通过在语言理解、知识准确度、逻辑推理等 8 个核心维度,利用 100+评测集对 LLM 进行综合评估. 根据其发布的 2025 年 3 月实时公开学术榜单 (使用学术开源评测集进行评测),深度求索公司研发的 DeepSeek-R1 模型表现领先 (榜单前 5 名数据如表 3 所示).

表 3 2025 年 3 月 CompassRank 大语言模型公开学术榜单

大语言模型名称	平均分数	是否开源
DeepSeek-R1	86.7	是
O3-mini-2025-01-31	85.6	否
O1-mini-2024-09-12	77.7	否
DeepSeek-V3	74.0	是
Doubao-1.5-pro-32k-250115	66.2	否

基于开源可复用性、技术性能及资源效率的综合考量,本研究采用本地化部署的 DeepSeek-R1:32B 模型作为核心需求生成引擎,模型温度参数 (temperature) 保留默认值 1.0 以平衡生成结果的多样性与可控性. 系统架构层面,基于 LangChain 框架实现多智能体间的任务编排与上下文感知式交互,降低流程复杂性.

(4) 评价指标设计. 为客观比较 BMHACT 框架与需求重用方法在创新需求生成效率方面的性能差异,本实验定义以下指标:(1) 迭代次数:达成有效生成 (CNU≥4.5) 时所需的迭代次数并设置最大迭代阈值为 10 次以避免无效迭代;(2) 时间成本:从需求输入到生成达标结果的全流程耗时 (单位: min). 通过构建“迭代次数-时间成本”二元评估矩阵,从而实现对方法效能的全面量化分析.

(5) 实验环境设置. 本实验在一台配备英特尔第 12 代酷睿处理器 (Intel Core i5-12600KF) 的计算机上运行,计算机的内存为 32 GB,图形处理器为 NVIDIA GeForce RTX 4070 SUPER,操作系统为 Windows 11 专业版. 实验采



用 Python 编程语言, 版本为 3.12, 开发和运行环境使用 PyCharm 集成开发环境 (IDE), 并利用 Anaconda3 进行环境管理与依赖配置, 以确保实验环境的隔离性与可重复性。

## 4.2 实验过程

### 4.2.1 BMHACT 框架的实施过程

在为 5 个智能体输入角色定义及中间件文档提示词后, 首先, 需求收集员智能体通过多轮对话建模用户需求, 构建结构化需求范围说明书, 明确系统边界与核心目标。随后, 业务流程分析员智能体基于系统目标锚定业务流程, 联网检索行业数据, 生成多维度对比矩阵 (流程效率、用户满意度、用户反馈), 识别优先级改进机会。在此基础上, 业务设计员基于矛盾分析框架与发散性思维策略生成候选创新方案, 并通过可行性-价值评估矩阵筛选输出创新业务需求表单, 从中随机抽取 3 条需求, 供后续质量验证阶段评估。

在质量验证阶段, 系统分析员对业务需求进行工程化处理, 分类功能和非功能需求并通过语言规范性检测模型 GRUEN<sup>[26]</sup>进行过滤, 得到需求候选池。随后, 领域专家与客户组成的评审委员会采用三维度五点量表进行评分, 结合对抗性验证机制 (专家质疑-LLM 反演响应) 迭代优化需求方案。最终, 通过评审的需求生成文档, 未通过项则启动反馈驱动的迭代优化, 直至满足终止条件 ( $CNU \geq 4.5$  或者最大迭代阈值 10)。

上述过程经过 5 轮迭代后达到终止条件, 最终轮各智能体输出的中间件文档如后文图 6 所示, 交付的最终需求文档见 <https://gitee.com/lxh8888/bmhact>。

### 4.2.2 基于对抗样本的追溯式需求生成方法实施过程

由于当前方法需基于原始需求进行扰动得到新需求, 为满足实证研究领域要求, 本文收集了三一重工、中联重科、徐工集团等同领域头部企业的 632 条门户网站需求文本。将中文需求专业翻译为英文后, 构建原始需求集  $R$  (标签“e”)。通过 NLTK 工具对  $R$  执行词性标注, 生成扰动候选词池, 并依据词频降序排序优先级。该操作同时为基于字符串匹配规则 (Levenshtein 距离阈值  $\delta=0.85$ ) 的扰动词选择提供基础支持。然后采用 5 种词级扰动策略 (Replace/Insert2L/Insert2R/Delete/ReplaceBi), 在原始需求集  $R$  上随机选择扰动位置, 快速生成与  $R$  等量的扰动需求集  $R_p$  (“n”标签), 与原始需求集  $R$  组成平衡数据集。

之后构建 GRU 分类器启动“扰动-分类”正反馈迭代优化, 每轮优化流程包含以下 3 步。

(1) 对抗样本识别: 筛选被误分类为“e”的  $R_p$  样本。

(2) 显著性驱动优化: 对每个误分类样本回溯其源需求  $r_j \in R$ , 计算其中每个词  $w_i$  的显著性得分:

$$Score(w_i) = \frac{1}{3} [F(w_1 \dots w_{i-1} w_i w_{i+1} \dots w_n) - F(w_1 \dots w_{i-1} w_{i+1} \dots w_n) + F(w_1 \dots w_i) - F(w_1 \dots w_{i-1}) + F(w_i \dots w_n) - F(w_{i+1} \dots w_n)], i = 1, \dots, n,$$

进而对最高得分词施加新扰动策略生成优化样本  $R_p'$ ;

(3) 训练集更新: 合并  $R_p'$  与未被误分类的原始  $R_p$  样本, 构成新一代  $R_p$ 。

为量化生成效果并控制流程, 每轮迭代输出误分类样本作为候选需求集, 并从中随机抽取 3 条进行双盲评估。评估环节由领域专家与客户组成独立评审组, 对抽样需求进行三维度量化评分, 以综合得分 CNU 作为核心指标, 当  $CNU \geq 4.5$  或迭代轮次  $\geq 10$  或对抗样本重叠率  $> 90\%$  时终止流程。

经 7 轮迭代后, 对抗样本重叠率达 92.3% ( $> 90\%$  阈值), 满足终止条件。最终交付的 3 条优化需求如后文表 4 所示。其中, 下划线表示在原始需求中扰动的实施词以及扰动后需求的对应改动。

### 4.2.3 基于需求重用的方法实施过程

首先, 从在线软件平台 Softpedia 抓取汽车资讯、企业官网及购物比价这 3 大领域共 1200 条软件功能描述。随后, 首先使用 Doc2Vec 模型<sup>[45]</sup>将每条需求转换为语义向量, 接着通过 BIRCH 算法<sup>[46]</sup>对向量化后的需求进行聚类, 最后基于轮廓系数 (Silhouette score)<sup>[47]</sup>确定最优聚类数。

在创新生成阶段, 实施跨簇元素组合策略。首先从非相邻聚类随机抽取动词、对象及修饰成分以生成跨领域的需求。随后通过 TF-IDF 余弦相似度衡量生成需求与现有需求的相似性, 选择相似度最低的 10% 作为候选。最终从低相似度候选集中随机抽取 3 条需求供人工评估。

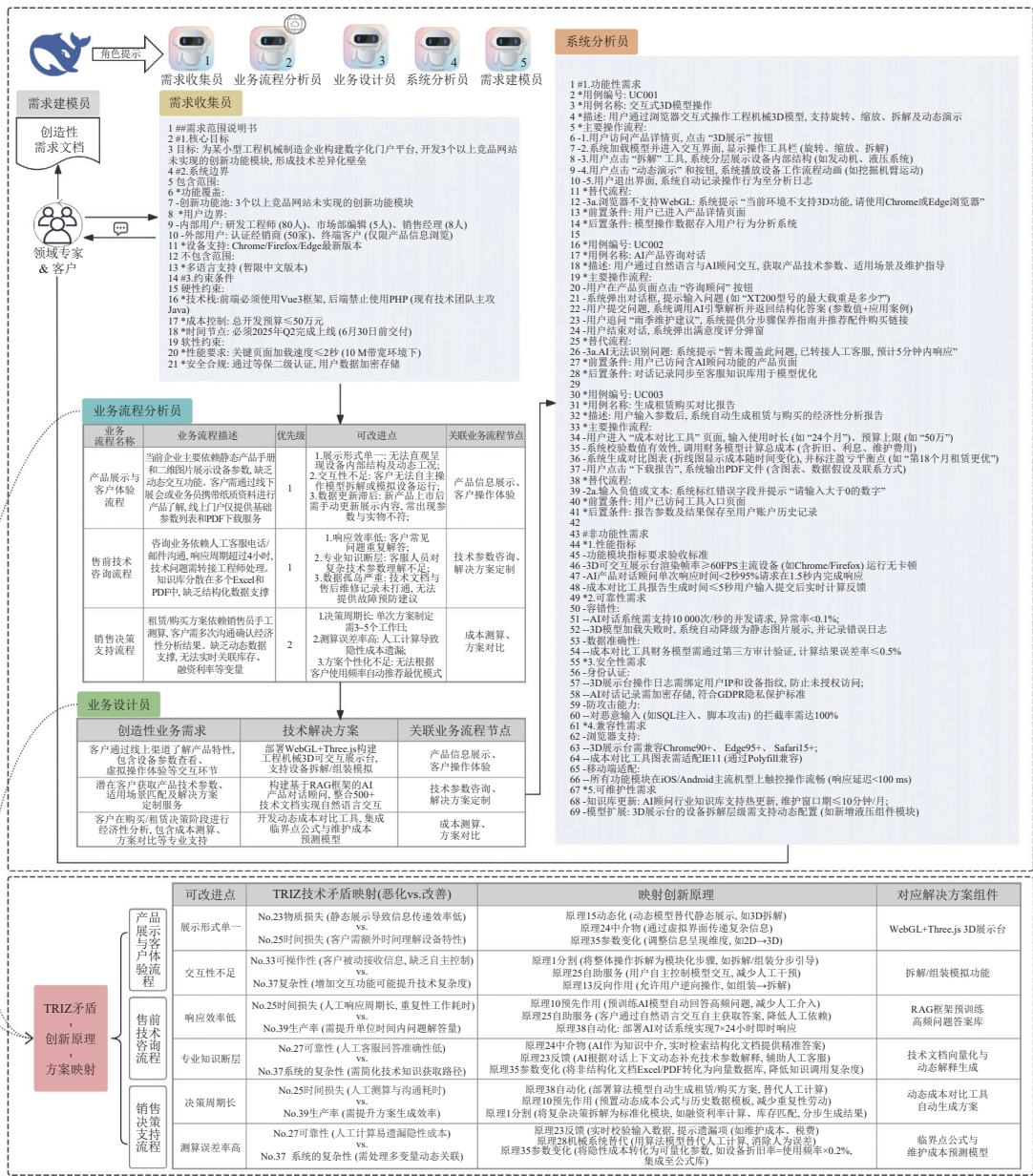


图 6 最终轮多智能体中间文档

表 4 基于需求重用方法生成的创新需求

需求编号	扰动策略	原始需求	扰动后需求
A-R1	Replace	The platform enables users to configure <u>products</u> and generate customized quote requests	The platform enables users to configure <u>logistics</u> and generate customized quote requests
A-R2	Insert2L	The system requires full <u>multilingual</u> support, enabling synchronous multilingual content editing and storage	The system requires full <u>containerized multilingual</u> support, enabling synchronous multilingual content editing and storage
A-R3	Replace	The system must associate technical parameters, <u>images</u> , videos, and other multimedia resources with each product/model	The system must associate technical parameters, <u>AR-models</u> , videos, and other multimedia resources with each product/model

与前面的实验相同, 评估环节由领域专家与客户组成双盲评审组, 对抽样需求进行三维度量化评分, 采用 CNU 作为迭代控制指标. 优化流程持续至  $CNU \geq 4.5$  或达到 10 次迭代阈值, 未达标需求返回跨簇重组环节.

上述过程经过 10 轮迭代, 触发迭代阈值后, 交付的 3 条需求如表 5 所示.

表 5 基于需求重用方法生成的创新需求

需求编号	需求描述
R-R1	The system shall provide user with ability to provide multi-angle product image display
R-R2	The system shall provide user with ability to implement a product specification comparison tool
R-R3	The system shall provide user with ability to integrate an online trial appointment function

4.3 实验结果与对比分析

4.3.1 实验结果

本研究通过对比实验验证了 BMHACT 框架相较于两组基线方法的优势. 如表 6 所示, 在生成等量需求的前提下, BMHACT 框架相较基线方法在迭代次数与总耗时两个核心指标上均呈现显著优势, 其中迭代次数分别降低 28.6% 和 50%, 耗时缩减 33.3% 和 66.7%. 实验结果有效回应了 Q1 问题, 证实本文框架在创新需求捕获效率方面具有明显提升.

表 6 实验结果

方法名称	迭代次数	总耗时 (min)	需求编号	需求描述
BMHACT框架	5	20	B-R1	功能点名称: 工程机械3D可交互展示台 描述: 用户可360°旋转、缩放、拆解设备模型, 查看内部结构及动态演示 价值: 提升产品透明度与用户信任度, 减少线下看样成本 实现方式: 集成Three.js框架, 基于WebGL实现浏览器端3D渲染
			B-R2	功能点名称: 产品对话顾问 AI智能体描述: 通过自然语言交互解答产品技术参数、适用场景及维护问题 价值: 7×24 h实时响应, 降低人工客服压力, 提高用户满意度 实现方式: 基于GPT-4微调行业知识库, 结合对话管理引擎 (Rasa)
			B-R3	功能点名称: 租赁与购买成本对比工具 描述: 输入使用时长、预算等参数, 系统自动生成租赁与购买的经济性分析报告 价值: 辅助用户快速决策, 推动高价值订单转化 实现方式: 内置财务计算模型, 结合用户输入生成可视化图表 (Chart.js)
基于对抗样本的追溯式需求生成方法 (基线A)	7	30	A-R1	The platform enables users to configure logistics and generate customized quote requests
			A-R2	The system requires full containerized multilingual support, enabling synchronous multilingual content editing and storage
			A-R3	The system must associate technical parameters, AR-models, videos, and other multimedia resources with each product/model
需求重用方法 (基线B)	10	60	R-R1	The system shall provide user with ability to provide multi-angle product image display
			R-R2	The system shall provide user with ability to implement a product specification comparison tool
			R-R3	The system shall provide user with ability to integrate an online trial appointment function

表 6 展示了 3 种方法生成的需求案例. BMHACT 框架通过结构化提示模板生成的需求 (B-R1–B-R3) 展现出更系统的创新特性. 其需求文档包含功能名称、描述、价值及实现方案这 4 个维度, 形成覆盖需求定义、可行性评估与技术落地的完整框架. B-R1 提出基于 WebGL 的工程机械 3D 交互展示方案, 通过技术创新提升 B2B 场景的采购透明度; B-R2 设计基于自然语言处理 (NLP) 的 AI 产品顾问系统, 重构传统客服的交互范式; B-R3 开发租赁-购买经济性分析工具, 融合运筹学模型赋能用户决策. 这些需求不仅突破传统组合创新范式, 更深度整合了行业前沿技术 (如交互式 3D、NLP、运筹优化), 展现出现实应用场景中的突破性创新潜力.

基线 A 生成的需求 (A-R1–A-R3) 则呈现出独特的语义变形特征. 该方法通过词级扰动技术在原始需求基础上生成创新变体. A-R1 将核心功能对象由“configure products” (配置产品) 替换为“configure logistics” (配置物流),

实现功能场景的实质性迁移; A-R2 通过在基础功能“multilingual support”(多语言支持)前插入“containerized”(容器化)技术限定词,在维持核心功能的同时扩展了系统架构维度; A-R3 则将关联资源类型从“images”(图片)升级为“AR-models”(AR 模型),推动产品数据维度实现静态媒介向交互技术的演进.这些需求在保持原始功能内核的同时,通过精准的语义扰动实现了功能边界创新,其创新机制体现为可追溯的功能变形,形成创新路径的透明化映射.然而该方法的创新性高度依赖分词数据集质量,其扰动词源局限于现有语料库的分词结果,在跨领域场景中面临语义断层风险.

相较前述两种方法,基线 B 通过跨需求簇元素重组形成 3 个需求(R-R1-R-R3),其创新路径呈现明显的组合特征.如图 7 所示, R-R1 和 R-R2 整合了购物比价需求簇中的多角度产品展示与规格对比功能, R-R3 则迁移了汽车资讯需求簇中的试驾预约模块.该方法在创新性维度存在双重局限: (1) 创新机制受限于既有需求知识库的边界约束,生成结果多呈现已有元素的线性组合特征,难以突破预设知识体系; (2) 基于 Rupp 样本<sup>[48]</sup>的模板设计导致需求文档信息密度不足,缺乏完整的需求描述框架.

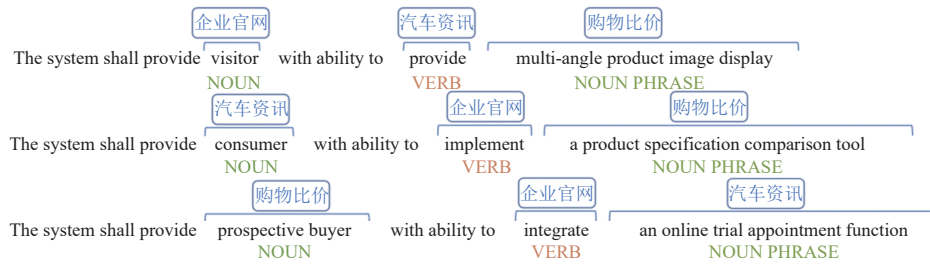


图 7 R-R1-R-R3 跨簇组合元素

#### 4.3.2 对比分析

本研究采用 Cohen's kappa 系数<sup>[49]</sup>评估人工评审组的评分者间信度,通过领域专家与客户代表构成的跨职能评审组( $n=6$ ),依据第 3.4 节设计的五点李克特量表对需求创新性进行多轮次评估.

##### (1) 评估共识度

图 8 展示了评审组逐轮对基线方法与 BMHACT 框架生成需求在清晰性、新颖性、实用性这 3 个维度的评分均值及误差.对每轮评分进行 Kappa 统计显示,基线 A 的需求评分 Kappa 值波动区间为 0.71–0.89 区间( $M=0.78$ ,  $SD=0.07$ ),呈现收敛态势;基线 B 的需求评分 Kappa 值波动区间为 0.65–0.83 ( $M=0.71$ ,  $SD=0.09$ ),呈现非稳态收敛;BMHACT 框架的需求评分 Kappa 值稳定于 0.72–0.93 区间( $M=0.85$ ,  $SD=0.06$ ),具有显著提升趋势.该结果与误差线长度变化呈现强相关性(Pearson's  $r=-0.89$ ),说明评估共识度随迭代过程持续而增强.

##### (2) 创新潜力指数(CNU)

我们使用第 3.4 节提出的 CNU 来度量需求的综合创新性.如图 9 所示, BMHACT 组首轮迭代即取得相对基线 B 的 CNU 优势值( $\Delta=0.17$ ),且在后续迭代中呈现单调递增趋势,于第 5 轮超过 CNU 达标值 4.5,3 个需求的 CNU 均值为 4.7 ( $SD=0.08$ ),满足迭代终止条件.相较而言,基线 A 组历经 7 轮迭代后 CNU 均值为 4.25 ( $SD=0.19$ ),未能达到 CNU 达标值,因对抗样本重叠率超过 90% 终止迭代.而基线 B 组历经 10 轮迭代后 CNU 均值仅为 3.83 ( $SD=0.08$ ),未能达到 CNU 达标值,因达到预设最大迭代次数限制而被迫终止.

创新潜力指数表现的显著差异,本质上源于 3 类方法创新机制的不同.基线 B 方法依赖历史需求的跨簇随机重组,其创新性受限于现有功能描述库的固有边界,本质上属于有限解空间的局部探索.基线 A 方法通过引入对原始需求的语义扰动,拓展了功能边界,但其创新性能在较大程度上受限于现有语料库中分词的质量.而 BMHACT 通过 TRIZ 矛盾矩阵的因果链分析与 LLM 跨领域知识迁移,构建了开放式创新空间,能够突破领域惯性思维的约束,生成更具创新性的需求.

##### (3) 创新性演进效率

本研究通过单因素方差分析来验证方法间创新性演进效率的差异性.针对 9 组需求样本的 CNU 分数进行组



间比较, 单因素方差分析显示 ( $F(8, 57)=4.294, p=4.33\times 10^{-4}$ ), 效应量  $\eta^2=0.3761$  (大效应). 该结果表明组间差异具有高度统计显著性 ( $p<0.001$ ), 且方法差异可解释 37.61% 的 CNU 分数变异. 根据 Cohen 准则, 该效应量超过大效应阈值 ( $\eta^2>0.14$ ), 从统计学角度充分支持 Q2, 证实 BMHACT 框架在需求创新性演进效率上的优越性.

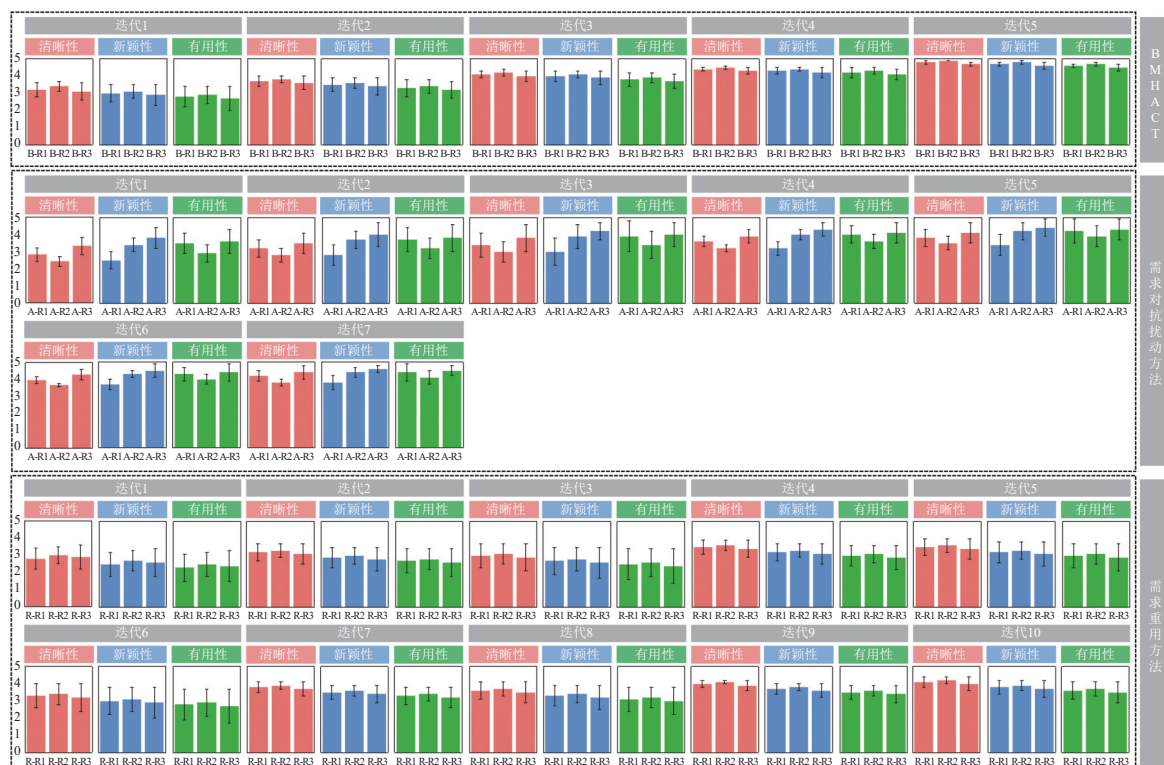


图8 逐轮次三维度评估数据

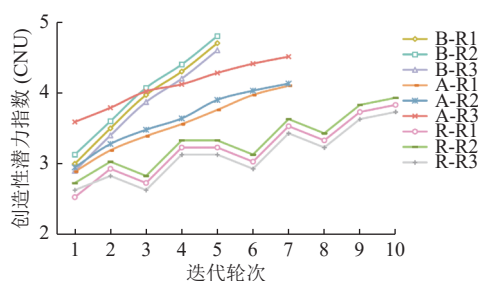


图9 逐轮次创新性潜力指数数据

我们认为这种效率差异源于方法论的本质区别. 需求重用方法依赖跨领域元素组合的随机探索策略, 缺乏系统性质量管控机制, 其需求演化过程呈现马尔可夫无记忆特性, 存在单次生成不可追溯修改的局限性. 基于对抗样本的追溯式需求生成方法则通过与分类器的协作优化, 可实现对需求文本扰动位置与扰动策略的迭代变更, 打破了单次生成不可追溯修改的局限, 但其仍缺乏系统性质量管控机制. 而 BMHACT 框架建立了多阶段质量门控机制, 具体表现为: (1) 在需求生成阶段, 通过预设文档模板强制结构化输出, 确保需求规格的完整性; (2) 在评估阶段, 构建强化学习引导机制, 将专家反馈转化为智能体团队的奖惩信号; (3) 在迭代优化阶段, 运用大语言模型的语义推理能力实现需求缺陷的因果追溯. 这种人机协同的闭环优化系统, 使需求 CNU 分数的平均增益率达到 0.34 每轮次, 较基线 A 和基线 B, 分别提升 2.16 倍和 2.81 倍.

4.4 需求创新性的进一步评估

4.4.1 评估设计

- (1) 评估方法. 为进一步验证框架的有效性并提升评估视角的全面性, 我们采用跨学科双盲评估设计, 从多维度验证本方法在需求创新性上的表现.
- (2) 志愿者招募策略. 这次评估我们通过分层抽样策略在相关院系 (软件工程系、机械制造工程系) 招募具有研究案例相关领域知识同源性的硕士研究生参与者 ( $n=30$ ). 根据年级维度 (研究生一年级至三年级) 与学科维度 (软件工程/机械制造工程) 构建  $3 \times 2$  抽样矩阵, 每个单元格招募 5 名志愿者, 确保样本在学术经验与专业背景维度均衡分布.
- (3) 调查问卷. 这份调查问卷分为两部分内容. 第 1 部分内容为表 6 中的 B-R1-B-R3、A-R1-A-R3 及 R-R1-R-R3 的需求描述. 第 2 部分是第 3.4 节中设计的五点李克特量表. 志愿者需要为每一项需求从清晰性、新颖性、实用性这 3 个维度进行打分, 每个维度对应 2 个问题, 并需要书面阐述打分理由.
- (4) 评估实施流程. 我们在测试开始前为志愿者提供了案例讲解与评分标准培训. 每位志愿者的平均评估耗时在 1 h 左右.

4.4.2 评估结果与分析

基于 30 份有效问卷的横截面数据, 我们开展了描述性统计与推论分析. 如图 10 所示, 相较于基线方法, BMHACT 框架在需求创新性的 3 个维度上均展现出显著优势, 这一结果与需求生成流程终止轮迭代的数据相吻合. 为进一步验证 6 人评审委员会与 30 人志愿者小组间的评分一致性, 我们针对终止轮迭代数据中 9 个需求各维度的评分, 以及本次 30 人志愿者对 9 个需求各维度的评分, 进行了不等方差双样本  $t$  检验. 结果显示, 各需求在各维度上的  $t$  值区间为 0.47~1.82 ( $p>0.05$ ), 表明两组评分具有较高的一致性, 为 Q2 提供进一步的统计支持, 证实 BMHACT 框架优于基线方法.

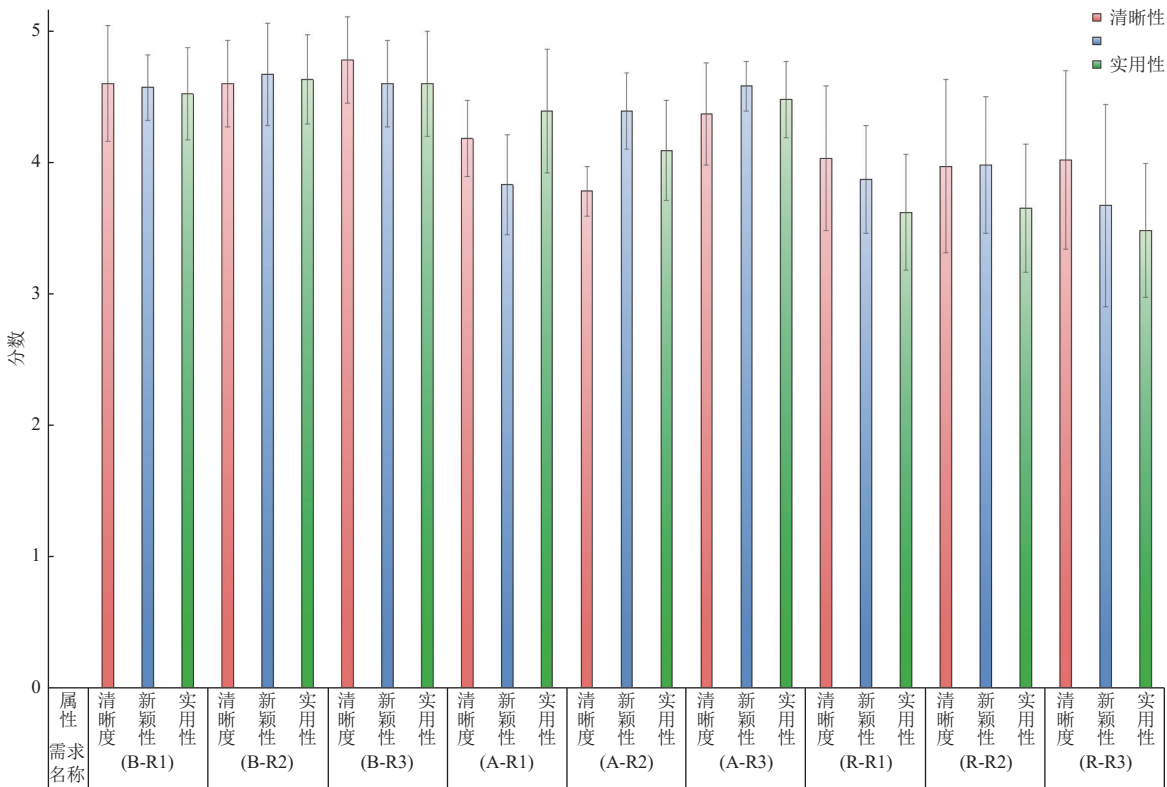


图 10 志愿者需求三维度评估数据

此外,通过跨学科评估视角的质性分析,我们还发现一些有趣的现象,不同专业背景的志愿者在需求特性维度呈现显著认知差异.以 BMHACT 框架生成的“工程机械 3D 可交互展示台”需求为例,采用 NLP 技术对两组志愿者(机械制造工程系  $n=15$ , 软件工程系  $n=15$ )的评分理由文本进行语义特征分析,并基于 Python WordCloud 库进行词汇分布可视化展示.如图 11(a)所示,机械制造工程系评分理由文本高频词汇集中于领域适配性维度(如 Solidworks Viewer、3D 模型、产品展示优化),其词频占比达 62.3%.而如图 11(b)所示,软件工程系则聚焦技术可行性维度(如 WebGL 渲染、Three.js 框架、高并发性能),相关技术术语出现频次占比 58.7%.



图 11 评分理由词云图

深入分析发现,来自机械工程的志愿者更关注需求的新颖性与产业价值,87%的高分评价( $CNU \geq 4$ )与“三维模型可视化可增强 B2B 采购决策”等产业应用论述相关.来自软件工程系的志愿者在认可需求创新性( $M=4.2$ ,  $SD=0.6$ )的同时,66.7%的评分理由涉及技术实现约束,特别强调需明确模型轻量化策略(如 LOD 技术应用)和分布式渲染架构设计.这种跨学科评估差异不仅充分验证了本方法生成需求的领域价值,也为我们下一步对智能体提示词工程的改进提供了一些新的思路.

## 5 总结与讨论

本文提出的 BMHACT 框架以 RUP 业务建模的人员构成与协作为蓝本,构建人-多智能体协作的创新需求捕获团队.通过 LLM 实时联网更新的知识库及其具备的强大语义理解与逻辑推理、跨领域知识迁移与发散性创意生成能力,设计若干中间文档提示模板,保障智能体产出物的可解释性与质量.该框架实现任务粒度逐渐降低的需求捕获流程,并引入人类专家进行生成需求的评估,形成“多智能体团队生成-人类验证反馈”的人机协作迭代闭环. BMHACT 框架的核心价值体现在: (1) 构建“多智能体团队生成-人类验证反馈”的迭代闭环,克服传统自动化方法存在的单次生成不可追溯修改的局限性; (2) 通过多智能体间的分工协作与 GRUEN 模型<sup>[26]</sup>自动化预筛,与传统工作坊方法相比,极大减少了人力资源投入; (3) 通过注入 TRIZ 理论的结构化提示词引导与 LLM 开放式创新的有机结合,达成需求创新性生成效率与质量的均衡优化.该框架具有较强的可扩展性,适用于不同行业和领域的中小型项目创新需求捕获任务,并且可以根据实际需求灵活替换不同的大语言模型,或对智能体提示词和产出中间件的提示词进行调整.在机械制造领域门户网站案例的实证研究中,证明了本框架在提升创新需求捕获效率和需求创新质量上的有效性.

BMHACT 框架在多智能体创新需求捕获团队的角色设置,主要参考了 RUP 业务建模中的人员构成和协作方式,这在一定程度上受到了传统需求工程方法论的影响,尚未充分考虑在智能化时代下创新需求捕获任务的新分工.在未来,我们将探索更多 LLM 之间的分工与协作机制和生成内容的质量保障模式,以更好地支持创新需求工程中的各种任务.此外,我们认为 BMHACT 框架提供的人机协作下的“多智能体团队生成-人类验证反馈”的迭代式方案,为软件工程中的其他任务提供了一个有效的探索思路,我们也计划将该框架应用于更多的软件工程任务中.

## References

- [1] Nuseibeh B, Easterbrook S. Requirements engineering: A roadmap. In: Proc. of the 2000 Conf. on the Future of Software Engineering. Limerick: ACM, 2000. 35–46. [doi: [10.1145/336512.336523](https://doi.org/10.1145/336512.336523)]
- [2] Nguyen-Duc A, Cabrero-Daniel B, Przybylek A, Arora C, Khanna D, Herda T, Rafiq U, Melegati J, Guerra E, Kemell KK, Saari M, Zhang ZY, Le H, Quan T, Abrahamsson P. Generative artificial intelligence for software engineering—A research agenda. arXiv:2310.18648, 2023.
- [3] Aaen I. Software innovation—values for a methodology. In: Proc. of the 4th Scandinavian Conf. on Information Systems. Oslo: Springer, 2013. 72–86. [doi: [10.1007/978-3-642-39832-2\\_5](https://doi.org/10.1007/978-3-642-39832-2_5)]
- [4] Ebert C, Abrahão S, Mani VS. From idea to impact: Survival guide for successful products. IEEE Software, 2024, 41(3): 20–25. [doi: [10.1109/MS.2024.3363540](https://doi.org/10.1109/MS.2024.3363540)]
- [5] Jing XL. Trends of innovation in software enterprises under the background of cloud computing: Based on technological and business model innovation. Journal of Shenyang Normal University (Social Science Edition), 2012, 36(6): 75–77 (in Chinese). [doi: [10.19496/j.cnki.sxxb.2012.06.022](https://doi.org/10.19496/j.cnki.sxxb.2012.06.022)]
- [6] Aldave A, Vara JM, Granada D, Marcos E. Leveraging creativity in requirements elicitation within agile software development: A systematic literature review. Journal of Systems and Software, 2019, 157: 110396. [doi: [10.1016/j.jss.2019.110396](https://doi.org/10.1016/j.jss.2019.110396)]
- [7] Maiden N, Robertson S. Integrating creativity into requirements processes: Experiences with an air traffic management system. In: Proc. of the 13th IEEE Int'l Conf. on Requirements Engineering. Paris: IEEE, 2005. 105–114. [doi: [10.1109/RE.2005.34](https://doi.org/10.1109/RE.2005.34)]
- [8] Maiden N, Gizikis A, Robertson S. Provoking creativity: Imagine what your requirements could be like. IEEE Software, 2004, 21(5): 68–75. [doi: [10.1109/MS.2004.1331305](https://doi.org/10.1109/MS.2004.1331305)]
- [9] Maiden N, Manning S, Robertson S, Greenwood J. Integrating creativity workshops into structured requirements processes. In: Proc. of the 5th Conf. on Designing Interactive Systems: Processes, Practices, Methods, and Techniques. Cambridge: ACM, 2004. 113–122. [doi: [10.1145/1013115.1013132](https://doi.org/10.1145/1013115.1013132)]
- [10] Jones S, Lynch P, Maiden N, Lindstaedt S. Use and influence of creative ideas and requirements for a work-integrated learning system. In: Proc. of the 16th IEEE Int'l Requirements Engineering Conf. Barcelona: IEEE, 2008. 289–294. [doi: [10.1109/RE.2008.54](https://doi.org/10.1109/RE.2008.54)]
- [11] Karlsen IK, Maiden N, Kerne A. Inventing requirements with creativity support tools. In: Proc. of the 15th Int'l Working Conf. on Requirements Engineering: Foundation for Software Quality. Amsterdam: Springer, 2009. 162–174. [doi: [10.1007/978-3-642-02050-6\\_14](https://doi.org/10.1007/978-3-642-02050-6_14)]
- [12] Sakhnini V, Mich L, Berry DM. The effectiveness of an optimized EPMcreate as a creativity enhancement technique for Web site requirements elicitation. Requirements Engineering, 2012, 17(3): 171–186. [doi: [10.1007/s00766-011-0133-0](https://doi.org/10.1007/s00766-011-0133-0)]
- [13] Ferrari A, Spoleini P, Donati B, Zowghi D, Gnesi S. Interview review: Detecting latent ambiguities to improve the requirements elicitation process. In: Proc. of the 25th IEEE Int'l Requirements Engineering Conf. Lisbon: IEEE, 2017. 400–405. [doi: [10.1109/RE.2017.15](https://doi.org/10.1109/RE.2017.15)]
- [14] Giunta B, Burnay C, Maiden N, Faulkner S. Creativity triggers: Extension and empirical evaluation of their effectiveness during requirements elicitation. Journal of Systems and Software, 2022, 191: 111365. [doi: [10.1016/j.jss.2022.111365](https://doi.org/10.1016/j.jss.2022.111365)]
- [15] Farfeleder S, Moser T, Krall A, Stålhane T, Omoronyia I, Zojer H. Ontology-driven guidance for requirements elicitation. In: Proc. of the 8th Extended Semantic Web Conf. Heraklion: Springer, 2011. 212–226. [doi: [10.1007/978-3-642-21064-8\\_15](https://doi.org/10.1007/978-3-642-21064-8_15)]
- [16] Bhowmik T, Niu N, Mahmoud A, Savolainen J. Automated support for combinational creativity in requirements engineering. In: Proc. of the 22nd IEEE Int'l Requirements Engineering Conf. Karlskrona: IEEE, 2014. 243–252. [doi: [10.1109/RE.2014.6912266](https://doi.org/10.1109/RE.2014.6912266)]
- [17] Maiden N, Zachos K, Petrianakis K, Lockerbie J, Chanpalangsri C, Ernst H, Kara S. Beyond just generative AI for discovering software opportunities. IEEE Software, 2025, 42(3): 41–49. [doi: [10.1109/MS.2025.3540672](https://doi.org/10.1109/MS.2025.3540672)]
- [18] Do QA, Bhowmik T. Automated generation of creative software requirements: A data-driven approach. In: Proc. of the 1st ACM SIGSOFT Int'l Workshop on Automated Specification Inference. Lake Buena: ACM, 2018. 9–12. [doi: [10.1145/3278177.3278180](https://doi.org/10.1145/3278177.3278180)]
- [19] Bhowmik T, Niu N, Savolainen J, Mahmoud A. Leveraging topic modeling and part-of-speech tagging to support combinational creativity in requirements engineering. Requirements Engineering, 2015, 20(3): 253–280. [doi: [10.1007/s00766-015-0226-2](https://doi.org/10.1007/s00766-015-0226-2)]
- [20] Do QA, Bhowmik T, Bradshaw GL. Capturing creative requirements via requirements reuse: A machine learning-based approach. Journal of Systems and Software, 2020, 170: 110730. [doi: [10.1016/j.jss.2020.110730](https://doi.org/10.1016/j.jss.2020.110730)]
- [21] Gudaparthi H, Niu N, Wang BY, Bhowmik T, Liu H, Zhang JZ, Savolainen J, Horton G, Crowe S, Scherz T, Haitz L. Prompting creative requirements via traceable and adversarial examples in deep learning. In: Proc. of the 31st IEEE Int'l Requirements Engineering Conf. Hannover: IEEE, 2023. 134–145. [doi: [10.1109/RE57278.2023.00022](https://doi.org/10.1109/RE57278.2023.00022)]
- [22] Boden MA. The Creative Mind: Myths and Mechanisms. London: Routledge, 2004.



- [23] Jin DM, Jin Z, Chen XH, Wang CH. ChatModeler: A human-machine collaborative and iterative requirements elicitation and modeling approach via large language models. *Journal of Computer Research and Development*, 2024, 61(2): 338–350 (in Chinese with English abstract). [doi: [10.7544/jssn1000-1239.202330746](https://doi.org/10.7544/jssn1000-1239.202330746)]
- [24] Kruchten P. *The Rational Unified Process: An Introduction*. Boston: Addison-Wesley Longman Publishing Co., Inc., 2003.
- [25] Michael O. *Inventive Thinking Through TRIZ*. Berlin: Springer, 2006.
- [26] Zhu WZ, Bhat S. GRUEN for evaluating linguistic quality of generated text. In: *Proc. of the 2020 Findings of the Association for Computational Linguistics*. Association for Computational Linguistics, 2020. 94–108. [doi: [10.18653/v1/2020.findings-emnlp.9](https://doi.org/10.18653/v1/2020.findings-emnlp.9)]
- [27] Guilford JP. *The Nature of Human Intelligence*. New York: McGraw-Hill Book Co., 1967.
- [28] Maiden N, Jones S, Karlens K, Neill R, Zachos K, Milne A. Requirements engineering as creative problem solving: A research agenda for idea finding. In: *Proc. of the 18th IEEE Int'l Requirements Engineering Conf*. Sydney: IEEE, 2010. 57–66. [doi: [10.1109/RE.2010.16](https://doi.org/10.1109/RE.2010.16)]
- [29] Murukannaiah PK, Ajmeri N, Singh MP. Acquiring creative requirements from the crowd: Understanding the influences of personality and creative potential in Crowd RE. In: *Proc. of the 24th IEEE Int'l Requirements Engineering Conf*. Beijing: IEEE, 2016. 176–185. [doi: [10.1109/RE.2016.68](https://doi.org/10.1109/RE.2016.68)]
- [30] Burnay C, Horkoff J, Maiden N. Stimulating stakeholders' imagination: New creativity triggers for eliciting novel requirements. In: *Proc. of the 24th IEEE Int'l Requirements Engineering Conf*. Beijing: IEEE, 2016. 36–45. [doi: [10.1109/RE.2016.36](https://doi.org/10.1109/RE.2016.36)]
- [31] Zachos K, Maiden N. Inventing requirements from software: An empirical investigation with web services. In: *Proc. of the 16th IEEE Int'l Requirements Engineering Conf*. Barcelona: IEEE, 2008. 145–154. [doi: [10.1109/RE.2008.39](https://doi.org/10.1109/RE.2008.39)]
- [32] Bhowmik T, Do AQ, Lam H, Amin R, Bradshaw GL, Niu N. On the way to a framework for evaluating creativity in requirements engineering. In: *Proc. of the 24th IEEE Int'l Conf. on Information Reuse and Integration for Data Science*. Bellevue: IEEE, 2023. 191–198. [doi: [10.1109/IRI58017.2023.00039](https://doi.org/10.1109/IRI58017.2023.00039)]
- [33] Xie XQ. *Research on RUP method and its application* [MS. Thesis]. Chongqing: Chongqing University, 2004 (in Chinese with English abstract). [doi: [10.7666/d.y704101](https://doi.org/10.7666/d.y704101)]
- [34] Nykänen O, Salonen J, Markkula M, Ranta P, Rokala M, Helminen M, Alarotu V, Nurmi J, Palonen T, Koskinen KT, Pohjolainen S. What do information reuse and automated processing require in engineering design? Semantic process. *Journal of Industrial Engineering and Management*, 2011, 4(4): 669–698. [doi: [10.3926/jiem.329](https://doi.org/10.3926/jiem.329)]
- [35] Decreus K, Poels G, El Kharbili M, Pulvermueller E. Policy-enabled goal-oriented requirements engineering for semantic Business Process Management. *Int'l Journal of Intelligent Systems*, 2010, 25(8): 784–812. [doi: [10.1002/int.20431](https://doi.org/10.1002/int.20431)]
- [36] Altshuller GS, Shapiro RB. Psychology of inventive creativity. *Issues of Psychology*, 1956, 6: 37–49.
- [37] Ilevbare IM, Probert D, Phaal R. A review of TRIZ, and its benefits and challenges in practice. *Technovation*, 2013, 33(2–3): 30–37. [doi: [10.1016/j.technovation.2012.11.003](https://doi.org/10.1016/j.technovation.2012.11.003)]
- [38] Silva GBL, Martinez GAS. Integrated TRIZ with QFD applied to automotive valve industry. *Holos*, 2018, 1: 50–60. [doi: [10.15628/holos.2018.6669](https://doi.org/10.15628/holos.2018.6669)]
- [39] Lee CH, Li L, Zhong P. TRIZ evolution trend-based public service innovation for enhancing social participation of life garbage classification. In: *Proc. of the 21st Int'l TRIZ Future Conf*. Bolzano: Springer, 2021. 57–73. [doi: [10.1007/978-3-030-86614-3\\_5](https://doi.org/10.1007/978-3-030-86614-3_5)]
- [40] Santis SHS, Marcicano JPP, Dedini FG, Held MSB, Sanches RA. The application of the classification of the theory of inventive problems solving (TRIZ) in textile design (2018). *Trends in Textile Engineering & Fashion Technology*, 2018, 2(2): 1–7. [doi: [10.31031/TTEFT.2018.02.000532](https://doi.org/10.31031/TTEFT.2018.02.000532)]
- [41] Zhang YY. Large language model in SD-WAN intelligent operations and maintenance. *Research Briefs on Information and Communication Technology Evolution*, 2023, 9: 178–188. [doi: [10.56801/rebict.v9i.177](https://doi.org/10.56801/rebict.v9i.177)]
- [42] Cai Z, Wang M, Zhao YY. Generative AI provides a new engine for the development of publishing. *All-media Explorations*, 2023, (5): 119–121 (in Chinese).
- [43] Tao JY, Xi XF, Sheng SL, Cui ZM, Zuo Y. Review on enhancing reasoning abilities of large language model through structured thinking prompts. *Computer Engineering and Applications*, 2025, 61(6): 64–83 (in Chinese with English abstract). [doi: [10.3778/j.issn.1002-8331.2405-0069](https://doi.org/10.3778/j.issn.1002-8331.2405-0069)]
- [44] Jin M, Zhou X. *Advanced Software Development Processes: Rational Unified Process, Agile Processes, and Microsoft Process*. Beijing: Tsinghua University Press, 2005 (in Chinese).
- [45] Le Q, Mikolov T. Distributed representations of sentences and documents. In: *Proc. of the 31st Int'l Conf. on Machine Learning*. Beijing: JMLR. org, 2014. 1188–1196.
- [46] Zhang T, Ramakrishnan R, Livny M. BIRCH: An efficient data clustering method for very large databases. *ACM Sigmod Record*, 1996, 25(2): 103–114. [doi: [10.1145/235968.233324](https://doi.org/10.1145/235968.233324)]

- [47] Rousseeuw PJ. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 1987, 20: 53–65. [doi: [10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)]
- [48] Pohl K, Rupp C. *Requirements Engineering Fundamentals: A Study Guide for the Certified Professional for Requirements Engineering Exam, Foundation Level, IREB Compliant*. Santa Barbara: Rocky Nook, 2015.
- [49] Savage T, Nayak A, Gallo R, Rangan E, Chen JH. Diagnostic reasoning prompts reveal the potential for large language model interpretability in medicine. *NPJ Digital Medicine*, 2024, 7(1): 20. [doi: [10.1038/s41746-024-01010-1](https://doi.org/10.1038/s41746-024-01010-1)]

#### 附中文参考文献

- [5] 景秀丽. 云计算背景下软件企业创新的趋势——基于技术创新和模式创新层面. *沈阳师范大学学报(社会科学版)*, 2012, 36(6): 75–77. [doi: [10.19496/j.cnki.sxxb.2012.06.022](https://doi.org/10.19496/j.cnki.sxxb.2012.06.022)]
- [23] 靳东明, 金芝, 陈小红, 王春晖. ChatModeler: 基于大语言模型的人机协作迭代式需求获取和建模方法. *计算机研究与发展*, 2024, 61(2): 338–350. [doi: [10.7544/issn1000-1239.202330746](https://doi.org/10.7544/issn1000-1239.202330746)]
- [33] 谢新泉. RUP 方法及其应用研究 [硕士学位论文]. 重庆: 重庆大学, 2004. [doi: [10.7666/d.y704101](https://doi.org/10.7666/d.y704101)]
- [42] 蔡喆, 王梅, 赵延永. 生成式 AI 为助力出版发展提供新引擎. *全媒体探索*, 2023, (5): 119–121.
- [43] 陶江垚, 奚雪峰, 盛胜利, 崔志明, 左严. 结构化思维提示增强大语言模型推理能力综述. *计算机工程与应用*, 2025, 61(6): 64–83. [doi: [10.3778/j.issn.1002-8331.2405-0069](https://doi.org/10.3778/j.issn.1002-8331.2405-0069)]
- [44] 金敏, 周翔. *高级软件开发过程: Rational 统一过程、敏捷过程与微软过程*. 北京: 清华大学出版社, 2005.

#### 作者简介

刘新华, 博士生, 主要研究领域为深度学习, 软件工程.

金敏, 博士, 教授, 博士生导师, CCF 专业会员, 主要研究领域为深度学习, 基于 AIGC 大模型与 NLP 的软件工程.

余梦蛟, 硕士生, 主要研究领域为深度学习, 智能化软件系统.

谢文涛, 博士生, 主要研究领域为深度学习, 药物副作用预测.