多智能体驱动新工科人才培养实验教学体系构建

王 婉 孔 梅
吉林建筑大学经济与管理学院，吉林 长春 130018

[摘要]在新工科建设持续推进与生成式人工智能加速融入高等教育的背景下，实验教学中的资源割裂、指导同质化和评价证据不足问题突显。本文运用文献分析与系统设计方法，以系统工程思想统筹教学要素，深入探讨多智能体技术与新工科实验教学的耦合机制。研究以 OBE 和 CDIO 理念为原则，提出由目标层、教学活动层、智能体协同层、学习证据层和治理保障层组成的总体架构。围绕实验教学的关键环节，重点设计实验资源动态关联组织与知识发现系统、多模态实验导学系统及人机协同智能评价系统，并提出实施策略，旨在形成可供后续教学的设计方案，为人工智能时代新工科实验教学改革提供理论参照与实践路径。

[关键词]多智能体；新工科；实验教学体系；大语言模型；人机协同；智能评价
DOI：10.33142/fme.v7i10.20705　　中图分类号：G642　　文献标识码：A

Construction of a Multi-agent-driven Experimental Teaching System for Talent Cultivation in Emerging Engineering Education

WANG Wan, KONG Mei
School of Economics and Management, Jilin Jianzhu University, Changchun, Jilin, 130018, China

Abstract: As Emerging Engineering Education continues to advance and generative artificial intelligence becomes increasingly integrated into higher education, the problems of fragmented resources, undifferentiated guidance, and insufficient evaluative evidence in experimental teaching have come to the fore. Employing literature analysis and system design methods, and drawing on systems engineering to coordinate instructional elements, this study investigates the coupling mechanisms between multi-agent technologies and experimental teaching in emerging engineering education. Guided by the principles of Outcome-Based Education (OBE) and the Conceive-Design-Implement-Operate (CDIO) approach, it proposes an overall architecture comprising five layers: objectives, teaching activities, agent collaboration, learning evidence, and governance and support. With a focus on the key stages of experimental teaching, the study designs a system for the dynamic association and organization of experimental resources and knowledge discovery, a multimodal guided-learning system for experiments, and a human-AI collaborative intelligent assessment system. Implementation strategies are also proposed. The study aims to develop a design scheme applicable to subsequent teaching practice, which providing a theoretical reference and practical pathways for reforming experimental teaching in emerging engineering education in the era of artificial intelligence.

Keywords: multi-agent systems; emerging engineering education; experimental teaching system; large language models; human-AI collaboration; intelligent assessment

1 问题提出

新工科建设回应的是产业技术体系、工程组织方式与人才能力结构的同步变化。自新工科研究与实践工作启动以来，相关政策不断强调学科交叉、产教融合和实践创新，工程人才培养由相对稳定的知识传授逐步转向面向新经济与复杂工程问题的综合能力塑造[1,2]。已有研究从新范式、专业结构与培养模式等角度阐释了这一转型，指出工程教育不能停留于课程内容增减，而应重新审视培养目标、课程体系与实践环境之间的关系[3-5]。实验教学处于知识理解与工程行动的交汇处，因而不是理论教学的附属验证环节，而是学生形成问题意识、设计能力和工程判断的重要场域。

大多数高校的实验教学活动一般按照一定的课程体系和相应的实验室来安排和执行，目的是保证教学内容能

Copyright © 2026 by authors and Viser Technology Pte. Ltd.

够准确地达到所设定的学术要求和实验条件。实验指导手册、设备操作指南、虚拟模拟素材和行业案例之间缺少语义上的联系，造成学生面对复杂的任务时不能很好地把需要使用的各种资源准确地匹配起来并加以整合。由于在有限的教学时间里，指导教师要面对的是来自各个专业的学生所提出的不同问题的需求，因此他们会采取一种比较通用的方法来解决这些问题，而不是针对每一个学生所存在的个性特点和具体需要做详细的分析。评价主要是对学生成交的实验报告、最终成果进行评定，而忽视了方案改进、故障排除、合作交流这些重要的环节所具有的价值，这就造成了学生综合素养得不到全方位发展的局面。该问题之间存在着密切的联系，资源管理策略的选择直接影响到学习路径的形成，进而决定可以获取到什么样的过程性数据。根据以上数据得出的评价结果又会反过来影响到资源分配以及任务分派的调整。单纯依靠某一时间段内用智能手段来完成工作，容易产生新的信息孤岛问题，而且会加重已经存在的低效工作流程的固化状况，成为制约整个系统效率和创新能力提高的重要因素。

大语言模型和多智能体技术给重新构建现有的关系赋予了可能，但是技术的可得性并不一定保证教学实践是有效的。通用问答系统可以产生答案，但是对于实验安全规程、教育目标等的理解以及对学生实际问题的认识可能会比较欠缺；自动化评分可以大大加快反馈的速度，但是如果标准设置不清楚或者存在模型偏差的话，就会给评价的有效性、准确性带来威胁。与此相反，本文研究的重点不是讨论是否采用人工智能的问题，而是要从设计角度出发，对三个相关联的设计难题进行探究和解决，即怎样恰当地把多智能体融入到实验教学中，不同的智能体怎样合作完成学习任务，教师怎样用透明、可解析的数据证据来保持自己在教学决策过程中占有的主导地位。因此本研究采用文献分析法和系统设计法相结合的方式，试图创建起一个包含教育理念、业务操作流程和技术职能主要方面在内的完整框架。经过上述分析之后得出详细的体系结构、主要模块以及实现的前提条件，为之后的基于学校情境的实际应用和实证研究打下了基础。

2 理论基础

2.1 系统工程视域下的 OBE-CDIO 整合

系统工程是从宏观目标的角度出发，找到组成系统的所有要素、各个部分之间接口和反馈关系，防止局部优化造成整体效果受到影响的一种方法。推进实验教学改革的时候，它的主要单元已经不再是单一的实验项目了，而成了包含着培养目标指向、精细规划的任务系列、合理调配的资源体系、助力学习的过程支持和产生结果的证据收集等各方面的教学生态系统。系统设计流程的第一步就是明确人才培养目标输出，其最终目的并不是仅仅完成平台的发布上线，而是依靠基于实证的学习成效评价来检验预定的目标是否得到有效地达成，并据此不断地改进和完善设计策略。

OBE 为这一系统规定方向。成果导向并非把指标简单分解到每份实验报告，而是先界定学生完成学习后能够表现出来的复杂能力，再反向设计任务、支持和评价，使预期学习产出、教学活动与评价标准相互一致[6]。CDIO 则为能力形成提供工程实践载体，将构思、设计、实现和运行组织为连续过程，使学生经历需求辨析、方案权衡、原型实现和测试改进[7]。二者结合后，OBE 回答“为何教、教到什么程度”，CDIO 回答“通过何种工程活动形成成果”，系统工程则负责检查各模块之间是否衔接。多智能体的功能配置必须服从这一逻辑，而不能以模型能力替代教学目标。

2.2 多智能体分工与分布式认知

多智能体系统是由许多有相对独立自主能力的智能体组成的，它们依靠各自的感知对象去认识周围的世界，完成一定的任务，并且依靠彼此之间的交流和合作来达成单个智能体无法单独完成的工作[8]。在实验教学中涉及资源检索、学情分析、操作指导、安全警报、效果反馈等诸多环节，把它们分成一个个独立的功能模块加以管理，既有利于确定任务范围，保证决策过程公开透明，又利于跟踪决策依据，采取精确的验证手段。此处所说的“分工”，并不是要把教师的传统职责分割开来交给自动化系统去执行，而是在于把那些有规律性、容易被索引、可以实现自动化的任务交给智能代理来辅助完成，但是涉及价值判断、突发事件处理、最终评价决定等重要的教育工作仍然要由教师亲自来做。

分布式认知理论进一步说明，认知活动并非只发生在个体头脑内部，还分布于工具、符号系统、同伴协作与环境之中[9]。据此观察智能实验教学，多智能体既是信息处理工具，也是学习活动中的认知中介。其合理作用在于呈现关键线索、连接已有经验、支持方案比较，使学生能够把注意力投入工程推理，而非代替学生完成推理。若系统直接给出操作步骤和报告结论，短期任务完成度或许提高，却可能削弱学生面对不确定问题时的独立思考。因此，导学智能体应遵循支架逐步撤除的原则：在陌生任务中提供必要提示，在能力增长后减少显性指导。支架理论所强调的，正是支持应位于学习者独立水平与潜在发展水平之间，

Copyright © 2026 by authors and Viser Technology Pte. Ltd.

并随能力变化而调整[10]。

2.3 “目标-资源-活动-评价”四维耦合

按照以上理论结构，本文主要从四个方面展开分析并综合评价多智能体系统同创新工程学科实验教学融合的过程。目标维度的任务就是把毕业要求和课程目标转化成可以衡量、可以观察的具体成果，目的是评价学生能不能正确识别出某种约束的情境，能不能恰当地分析出异常数据的表现形式，能不能用逻辑推理来有效地阐述自己的选择方案的优点和缺点。经由创建资源维度同目标和任务间的语义联系，从而达成对设备、实例、知识体系以及标准规程等的实时调整，进而精确契合某个学习目的。就活动维度而言，将智能化辅助深入到以创新为灵魂的教学模式以及工业协同（CDIO）体系当中去，在各个阶段给予递进式的认知引导，既明显改善了学习者的感受，又很好地推动了他们实践技能的发展与提高。评估框架把过程性、结果性证据结合起来，用量化的方法来衡量目标达成的程度，用反馈循环把诊断的信息反向传送到任务规划和资源配置环节中去，促使后续的策略改进和资源调配不断向前推进。

目标解析、评价标准、风险等级、干预阈值这四个主要方面一起形成了一个以数据为依托的教学决策链条。该流程链路创建和改进的过程是由教育领域专业人士共同讨论决定的。在此结构之下，智能实体依照赋予它的权限去展开搜索工作，给出建议，整合数据，并且发出警报。系统特意塑造出一个人工修正、纠纷上诉和过程追查的功能模块，目的是创建起一种责任分明、人机协作的工作环境。

3 实验教学体系的总体设计

3.1 构建原则

系统架构的创建要保证产出同实证相契合。对每一个实验任务都要有明确的目标，即学习成果、学生需要展示的作品、评价的方式等，防止由于目标太宽泛而造成后续的评价只看作业的长短和正确答案。工程流程完整性要得到保证。把基本技能训练融入到有实际意义的情境当中，并且在综合性项目里包含对于需求变化、性能约束和策略改进等方面的考虑，从而让学习者充分体会工程决定背后的种种状况及权衡取舍。我们应该提倡并践行人机协作，在这种协作的过程中合理地运用人工智能技术。系统参与和介入的程度要按照教育需求以及风险级别来确定。处理高压电力应用、操作危险设备、开展伦理评价、评判学生成绩这些重要事情的时候，不能仅仅依靠模型预测的结果来进行自动化的决定，而应该小心谨慎地对待它可能存在的风险和复杂情况。最终目的就是建立一个开放互联的系统，坚持最少数据原则。要求资源可以共享、整合，身份验证机制、数据接入点要保持一致，只获取达到教育目的所需要的信息，不能因为全面考评而无限扩大监控范围。

3.2 五层总体架构

目标层处在系统的最高层次上，它由专业培养目标、课程学习成果和实验任务指标三者组成，它们之间联系用建立起来的映射矩阵来表示清楚。在指标定义的时候应该尽量做到准确并且有操作性，先选择可以直接观察到的行为动词来表达目标。同时要明确规定执行任务所处的情境和条件，给出评价任务完成情况质量标准。就创新意识而言，不能只是强调创新意识，而应该让学生根据一定的条件提出至少两个方案，并对方案的安全性、经济性、可行性进行详细的分析，从而确定出一个最佳的选择。这样的表述给教师提供了一个评判的标准，也给智能系统识别和检验证据设定了一个清晰的范围。

教学活动层按照基础训练、综合应用、创新探究的顺序来构建任务体系，用 CDIO（创意、设计、实现、运作）全过程贯穿于课前准备、课堂实施和课后延伸之中，从而达到全方位提升学生实践能力和创新能力的目的。基础训练主要是规范的操作流程和基本的概念理解，综合应用部分是多个模块知识的综合运用以及团队合作的任务安排，创新探究阶段让学生自己提出问题并思考解决办法，从而培养学生的创造思维和动手能力。三个层次不是固定的、特定的学年对应的，而是根据学生以前的知识、经验来灵活地配置和调节。智能体协同层创建起资源管理、引导学习、评价助力、流程调配、安全保障等种种重要角色，这些角色之间依照标准的任务清单开展信息交流，只传送必需的数据，不会出现无谓的信息共享情况。

学习证据层担负起把预习笔记、实施方案的各种版本、操作流程日志、实验得到的数据集、讨论会要点、成果报告、后续反思总结这些材料综合起来的任务。该层次主要针对每一个证据给出它的原始出处、产生的时间、对应的学生活动阶段做出标注。它不但是学生获得反馈的一种方式，也是教师对评价标准进行反思、检验的过程。治理保障层是整个体系架构中不可缺少的部分，它的主要功能有教师终审流程、权限管理、模型和知识库版本控制、内容安全保护、个人信息保护、异常情况下系统退出等，保证了系统可以正常运转并且数据安全。五层架构形成了一种体系化的模式，把教学目标的确定、精心安排的学习活动、智能辅助手段的应用、证据的搜集与分析、治理框架的创建有机地联系在一起，构成一个环环相扣、有很强可追溯性的链条。这样一种设计使得技术推动的教学实践活动一

Copyright © 2026 by authors and Viser Technology Pte. Ltd.

直处在教育制度的约束之下进行，进而达成技术和教育理念相融合并不断受到监管的目的。

3.3 构建与迭代流程

需求分析阶段不宜从采购平台开始，而应先梳理课程图谱、实验条件和学生学习困难。专业团队可通过毕业生访谈、企业任务分析、课程资料审阅和课堂观察，识别高价值但支持不足的实验环节，再判断哪些问题适合智能技术介入。对于主要源于设备数量不足或任务设计失当的问题，应优先改善基础条件，而非以技术替代必要投入。

体系设计阶段形成目标、任务、资源、证据对照表，明确各层次实验项目之间的先修关系与综合接口，并为每项智能功能指定输入、输出、责任人和禁用边界。智能体嵌入阶段采取模块化开发：先建设经教师审核的实验知识库和规则库，再接入检索、导学及评价服务；每个模块均须通过典型任务、边界任务与错误案例测试。试点评价阶段同步考察学习表现、使用体验、教师工作变化和系统风险，不能只用点击量或满意度证明教学价值。迭代优化阶段依据证据调整任务和支架，更新失效资源，记录模型版本变化及其对评价结果的影响，形成可复盘的质量改进过程。

4 三类核心系统及其协同运行

4.1 实验资源动态关联组织与知识发现系统

资源系统的基础不是简单汇总文件，而是建立统一、可维护的实验资源语义模型。资源采集智能体对实验教材、设备手册、安全规范、虚拟仿真项目、科研案例和企业任务进行分类与格式处理，保留作者、来源、适用课程和版本等元数据；知识组织智能体在教师审定的本体框架下抽取实验项目、知识点、能力指标、设备、风险、工程场景等关系；资源服务智能体依据任务阶段、学生已具备的先修知识和设备可用状态生成候选资源组合。涉及安全规范与操作参数的内容只能来自经审核来源，不能由模型自由补写。

动态关联性是指资源链接会因为任务特性以及证据基础的改变而不断更新、调整。就多学生在某一实验阶段经常出现概念混淆的情况而言，教育技术系统应当创建起一种机制，这种机制可以找出那些学生的错误表现同教师给予的反馈注释和相关学习资料之间的联系，进而给之后的教学引导赋予有针对意义的支持。同时设备操作指南一旦修改，就要把旧版的手册和所牵涉到的任务加以标注，然后加入审核当中，从而保证所有的相关教学活动是及时有效的。知识发现的重点不是去寻找没有被发现过的联系，而是帮助教育工作者找到并认识课程体系里的连贯、整合之处。课程规划里，机械结构设计任务同材料性能检测、制造成本核算这些要素融合起来，就构成了一个跨学科项目。是否将其纳入教学大纲要根据该课程教师对它的前置知识要求和课程时间安排来决定。

推荐结果应附带理由与来源，说明为何在此时此刻推荐此项资源。系统可提供不同难度的路径，而不是以单一最优路径替代学生选择。对于检索增强生成，可将经过审核的实验知识库作为外部证据来源，使模型回答与具体文档片段相连接；检索增强生成技术为知识密集型任务提供了结合参数化模型与外部检索的思路，但在实验教学中仍需配合来源白名单、版本控制和人工抽检，不能把检索到文本等同于结论正确。

4.2 基于多模态大模型的实验导学系统

实验导学系统用情景化支持的方式，与学生目前正在进行的学习任务相联系，给学生提供个性化的、有针对性的教学辅助。多模态感知组件用来解析学生主动给出的文本、图片、视频、音频的问题和实验数据，找出里面包含的对象以及它们的状态。任务理解模块就是把各种各样的输入同具体的 CDIO 周期阶段、既定实验目的和安全准则联系起来，保证教育活动是连续的、合法的。根据实际情况选用提问引导、概念讲解、过程指导、参照案例呈现或者教师介入等方式来满足不同的学习者需要和情境。交互界面用对话的形式、图表注释或者视觉对比等方式，直接给学生提出建议，使学生能够理解并掌握知识。对于使用持续性高侵扰性摄像头监控技术，应该非常慎重地对待，在没有其他方法可以达到目的的时候才可以考虑将其作为全天候行为监督的方法。目的是平衡教育环境的隐私保护和有效的管理之间关系，在做出这样的决定之前要对它是否有必要、可能产生的影响进行充分的考虑。

导学策略应体现先诊断、后提示。学生初次接触复杂装置时，系统可通过关键问题帮助其核对实验条件和操作顺序；出现错误时，优先引导学生比较预期现象与实际现象、定位可能变量，而非立即给出完整答案；当学生已经能够独立完成后，系统逐步减少过程提示，转向开放性追问和迁移任务。已有大语言模型多智能体导学研究尝试将目标映射、技能差距识别、学习路径规划和内容生成分工组织，为角色协同提供了可参考的技术方案。不过，具体实验课程仍须根据安全等级与教学目标重新设计提示策略，不能直接移植通用导学流程。

多模态模型产生出来的结果带有概率性质，视觉识别的过程也会受到光线状况、角度改变、设备不同这些因素的影响。为了保证系统的准确性以及决策的可靠程度，它应该可以清楚地区分出三种信息处理方式，即直接观察得

Copyright © 2026 by authors and Viser Technology Pte. Ltd.

到的事实、按照既定规则推导出的结果和还需要加以证实的意见。当出现图像置信度低、数据超出正常工作范围或者引起安全联锁的时候，教学辅助系统要停止给学生提供自动化的帮助，并且要向教师或者实验室技术人员提出求助。所有的核心指示都要保留对应的任务标识、证据来源和模型迭代编号，在之后的评价当中才能找到错误产生的原因，是由于信息感知阶段、知识获取阶段还是策略选择阶段造成的。

4.3 基于大语言模型的人机协同智能评价系统

智能评价的主要目的不是把实验结果变成一个可以自动产生的分数，而是在很大程度上提高证据归纳、反馈诊断的效果和准确性。在对系统的设计进行评价的时候，应该把量规当作中心点，仔细地把它的实验目的分成问题定义、设计方案、执行规程、数据分析、成果解读、合作与自我反思等诸多方面，并且对每一个方面的特点在各个层次上加以具体的表述。对于可以被规则引擎验证的客观任务，比如设备连接状态确认、数据完整性检查、代码测试结果判定、保证报告里所有字段齐全等等，都可以用规定好的规则去检测是否正确无误、符合要求。但是需要对语义有深刻的理解的任务，利用大型语言模型可以帮助找到主要观点并检验证据是否有效，也可以帮助判断结论是否合理。就包含创新价值的评价、工程伦理考虑、复杂的解决办法的选择这些高危决策而言，教师的作用显得十分突出，必须依靠完备的专业知识和阅历来作出最后的判定并给予引领。

过程证据的采集应服务于量规，而不是追求数据数量。预习文本可以反映概念准备，方案版本能够呈现设计变化，故障排查记录可揭示推理路径，团队讨论摘要则有助于理解决策过程。学习分析研究强调利用学习过程数据支持理解与决策，但数据进入评价前仍需验证其真实性、代表性与可比性。尤其不能把登录次数、对话长度等易获得指标直接当作学习投入，更不能因学生使用不同的表达方式而形成隐性惩罚。

收集过程证据应该成为衡量标准的主要依据，而不是仅仅为了达到数量上的积累。预读材料可以体现出概念准备情况，设计方案的各种版本显示出了设计发展的过程，故障诊断日志表现出了问题解决的思路，团队会议总结给人们提供了一种认识决策机制的方式。学习分析研究主要是利用学习过程中产生的数据去理解并指导决策，因此保证数据真实、有代表性、可比较就成为前置工作。特别地，在评价学习投入的时候，不应该把容易获得的指标，比如登录次数或者对话时间等直接当作实际的学习参与度，也不能因为学生所使用的沟通方式不同而无意中造成隐性的评价压力。

4.4 多智能体协同

过程协调智能体以实验任务单为工作对象，将目标、阶段、学生授权状态和资源状态转化为必要上下文，再调用相应服务。例如，学生提交异常数据后，导学智能体先判断是否需要概念提示，资源智能体提供经审核的数据处理材料，评价智能体只记录该事件及其后续修正，不把一次错误直接固化为能力缺陷。各智能体传递的是完成当前任务所需的信息，而非不受限制的个人画像。

当输出发生冲突时，系统按照安全规则优先、课程标准其次、教师指令再次、模型建议最后的顺序处理，并保留冲突记录。对于资源推荐与导学策略不一致、规则评分与模型评分差异过大等情况，应自动降级为人工审查。所谓协同，不是让更多模型反复讨论直至形成表面一致，而是通过明确接口、证据校验和责任归属减少错误传播。系统还应设置独立审计角色，定期检查知识来源、提示日志、异常中断和人工修改，防止某一智能体的错误在链路中被逐层放大。

5 实施策略

5.1 以教学责任为主线的组织保障

多智能体实验教学组织设计时要重视整合教务管理、专业学院、实验中心、信息化部门这些方面的资源，并且要保证同校企合作单位之间有良好的配合，以免产生技术部门独自创建系统、教学部门无法使用的情况。从学校角度来说，应该创建起一个以教学负责人为核心的工作体系，主要工作就是对教育标准进行审核，把各种教育资源融合起来，并且开展风险管理工作。专业学院主要是把总目标分解成具体的任务，完成项目的设计与实施，对学生成绩进行评价。实验中心的任务就是保证实验设备的兼容性、操作过程的安全性并制订出相应的应急预案。信息化部门给予重要的技术支持，即学生身份验证、系统日常运行维护、数据安全保障等。企业的代表参加进来，用实际的工作事例以及审核行业的标准等方式来充实教学内容，但是不干预对学生个人评价的过程。

5.2 模块化技术支撑与全周期数据治理

模型选择应服从任务需求：规则明确的工作不必调用高成本生成模型，涉及本地实验资料的敏感任务可优先评估校内部署或受控服务。知识库、提示模板、评价规则和模型版本均需独立管理，升级前使用固定案例集开展回归测试，防止同一表现在不同时间获得不可解释的差异评价。此外，数据治理贯穿采集、传输、使用、保存与删除全过

Copyright © 2026 by authors and Viser Technology Pte. Ltd.

程。研究使用教学数据时，还应另行履行伦理审查与授权程序。大语言模型教育应用研究已提示隐私、偏见、透明度与过度依赖等实践和伦理问题[11]；生成模型的幻觉风险也意味着流畅回答不能作为事实正确性的依据[12]。因此，知识检索、规则校验、人工复核和异常退出不是附加功能，而是系统投入教学使用的前置条件。

5.3 面向人机协同教学的教师发展

教师培训不应停留在提示词技巧和工具演示，而应围绕真实实验任务建立设计能力。其一，学习把学习产出转化为可观察证据，辨别哪些表现适合机器辅助判断，哪些必须由教师评价；其二，理解检索增强、模型置信度和多智能体接口的基本逻辑，能够追问系统建议来自何处；其三，通过错误案例研讨识别幻觉、偏差和自动化依赖，掌握降级与人工接管流程；其四，依据学习证据调整实验任务，而不是被平台预设指标牵引。教师角色将更多转向学习设计者、过程诊断者和质量责任者。人工智能教育研究早期综述曾指出，技术应用讨论中教师视角相对不足，这一提醒对实验教学尤为重要。

5.4 以学习证据和效度检验为核心的质量监控

教育质量监督体系应该覆盖教学过程和管理系统的各个方面，使每一个环节都达到高标准，既包括课堂教学实践，又包含技术平台应用，从而保证教育活动质量、效率。从教育学的角度来说，评价重点应该包含目标达成情况、技能提升幅度、作业难易程度是否合适、反馈是否及时有效等各方面。就系统层面而言，要对它的搜索精度、采纳改进措施之后的效果、评价标准的一致性、异常中断的原因以及背后是否需要人工干预的合理性展开全面考虑并作出评判。经过对课程实践的观察发现有两个重要的交集，一个是系统使用频繁的时候，如果学生的解决问题能力没有得到应有的增长，那么就说明任务设置和辅助工具的设计存在缺陷，需要进一步探究它们的匹配度以及有效性；另一个是自动化评价同人工评价出现较大差异的时候，单纯依靠加大模型调用次数来试图缩小差距的做法并不是长久之计。相反，应该立即停止目前的评价工作，开展一次全面的评价标准校准工程，以保证两种评价方式可以达到很高的一致性，并且准确地反映出真实的学生成果，从而推动整个评价体系的改善和公正。后续研究可使用准实验设计、交叉对照设计或者分阶段实施设计的方法论，在同一个学科范围内，对相同的教育环境和学习者群体进行系统的比较和分析，从而得出不同的支持策略效果及影响。评估流程不单只是传统的知识测验和操作技能评定，还要加入迁移任务，目的在于全方位检测学生在陌生环境下自主应对繁杂问题的能力和适应情况。通过对学生的深入访谈可以全面了解学生面对困难时所采取的具体方法以及思考的过程，进而获得更加深刻的体会。

6 讨论

6.1 体系设计的教学价值

本文的主要贡献就是创建起一个透明化的体系，把目标设定、资源分配、执行过程、评价指标等各方面有机地联系起来，形成一条有条理、可追踪的链条，而不是简单地把三个智能模块拼接在一起。资源管理系统按照不同的任务要求以及技能提升的目标来重新组合实验材料，用逐渐减少的指导干预来模拟工程实践的教学方法，可以有效地促使学生展开自主探究并提升学习能力。将方案优化循环同问题解决策略融合起来，把它们当作决策架构的关键部分，这样可以保证学习过程是高效的并且具有很强的针对性。将任务单和治理规则结合起来使用之后，教育者可以准确找到学生在学习过程中遇到的重要问题，而且该种设置使学生清楚知道自己的反馈来自哪里、原因是什么，从而主动参与到自我发展的过程之中来，达成个性化的成长与改进的目的。智能技术的作用也发生了很大的变化，它已经从原来单纯的解决某个问题的工具变成了一个按照一定的教育规律和原则来给学生提供最基本的支撑和引导的综合性系统。

该结构既确定了跨学科合作的概念，又具体地提出了开展此类合作所要采取的措施和途径。新兴工科整合项目实施过程中会遇到很多难题，比如课程目标不同、资源分配失衡、教师之间沟通慢等都会影响到项目的持续发展能力，给它的长久运行造成较大障碍。保证一致性目标达成、清楚任务定义并选择合适的证据映射方法可以很好地发现知识上的欠缺之处，进而找出存在差异的地方，推动问题解决和决策改进的过程。在这种情况下，多智能体系统被用来支持信息检索和阶段性成果的整合工作，使教育领域内的专业人员可以把更多的精力放在创新性的任务规划以及重要的工程决策上，从而明显提高整个教育过程中的效率和质量。

6.2 研究局限与后续方向

本文主要研究的是概念建构、系统设计方面的内容，目前还没有形成一个完整的基于某个专业领域持续教育数据的评价体系来检验它的有效性，所以在这一阶段不对其对学生学业成绩的影响、教师工作量的变化和评分准确性的提高做出预测或者结论。所提出的智能体角色是根据逻辑功能进行划分的，并不是必须按照相同的数量去部署独立的模型。由于各个工科专业安全标准、设备配置和产

Copyright © 2026 by authors and Viser Technology Pte. Ltd.

出形式不同，在将架构具体化的时候需要对它进行再实例化，以满足各自的需求。

后续研究可以针对某一个实验课程设计展开系统的探究，重点在于对任务设置、支持手段和评价准则等要素在不断改进过程中所发生的变化进行详细记载。经过以上一系列操作之后，可以对这些元素怎样随时间变化进行了解，并且借助跨学科比较来评判目前存在的各种框架，在不同的环境里是否合适有效。从技术上来说，我们主要关注多模态证据的可靠度检验、解释评价的可追溯性分析、多智能体系统中异常传播机理的研究；在教育方面要对学生的自我调节能力、团队协作意识、独立完成工程任务的能力等进行考察，从而保证这些重要的能力得到良好的培养和发展。为了让多智能体技术真正成为新工科人才培养的助力而不是新的形式主义束缚，就必须把系统运行的关键性能指标放在首位，这样才能达到它的最大价值。

7 结论

就新工科实验教学中资源供给、教学指导和效果评价三者互相依赖又互相影响的难题来说，本文用系统工程理论同成果导向教育（OBE）的主要产出逻辑相结合，并且联系到工程实践环节，创建起一个包含全部要素且完备的实验教学架构。该框架把目标导向、资源整合、活动执行和成效评价这四个方面紧密联系起来，用多层次的结构来改善教学效果。本文提出了一种以目标设定层、教学实践活动层、智能主体协作层、学习成果证据层和治理体系保障层为构成的实验教学体系，目的是使教学目标更加准确地确定下来，教学过程更加有效地进行管理，学生的能力得到全方位的发展，并且教学效果可以客观地被评价出来，从而促进新工科实验教学的现代化、个性化发展。

根据以上结构，本文建立三个大的系统即资源动态关联和知识挖掘系统、多模态实验指导系统、人机协作智能评估系统。这些系统依靠任务清单、证据链条以及冲突调解机制的整合运用，达成高效协同运作的目的。体系实施的关键不在于模型数量的堆积，而在于保证教学目标的界定准确无误，知识来源的可信度足够支持有效的学习，教师的角色和职责清楚明确，质量提升策略依靠实证研究不断推行并且取得明显成效。为了对它的学习效果、适用范围和治理成本进行检验，后续的工作应该有基于学校具体情况的案例研究和跨学科实证分析，从而进一步完善这个课题。

基金项目：2025 年度吉林省教育科学规划一般课题：《多智能体驱动的新工科人才培养实验教学体系改革与实践》（GH25094）。

[参考文献]

[1]教育部高等教育司.关于开展新工科研究与实践的通知[Z].教高司函〔2017〕6 号,2017.

[2]教育部,工业和信息化部,中国工程院.关于加快建设发展新工科实施卓越工程师教育培养计划 2.0 的意见[EB/OL].(2018-09-17)[2026-11-11].http://www.moe.gov.cn/srcsite/A08/moe_742/s3860/201810/t20181017_351890.html

[3]吴爱华,侯永峰,杨秋波,等.加快发展和建设新工科 主动适应和引领新经济[J].高等工程教育研究,2017(1):1-9.

[4]顾佩华.新工科与新范式:概念、框架和实施路径[J].高等工程教育研究,2017(6):1-13.

[5]林健.面向未来的中国新工科建设[J].清华大学教育研究,2017,38(2):26-35.

[6]顾佩华,胡文龙,林鹏,等.基于“学习产出”（OBE）的工程教育模式——汕头大学的实践与探索[J].高等工程教育研究,2014(1):27-37.

[7] CRAWLEY E F, MALMQVIST J, ÖSTLUND S, et al. Rethinking Engineering Education: The CDIO Approach[M]. 2nd ed. Cham: Springer,2014.

[8] WOOLDRIDGE M. An Introduction to MultiAgent Systems[M]. 2nd ed. Chichester: John Wiley & Sons,2009.

[9] HUTCHINS E. Cognition in the Wild[M]. Cambridge, MA: MIT Press,1995.

[10] WOOD D, BRUNER J S, ROSS G. The role of tutoring in problem solving[J]. Journal of Child Psychology and Psychiatry,1976,17(2):89-100.

[11] YAN L, SHA L, ZHAO L, et al. Practical and ethical challenges of large language models in education: A systematic scoping review[J]. British Journal of Educational Technology,2024,55(1):90-112.

[12] JI Z, LEE N, FRIESKE R, et al. Survey of hallucination in natural language generation[J]. ACM Computing Surveys, 2023,55(12):1-38.

作者简介：王婉（1983—），女，汉族，吉林长春人，博士，副教授，吉林建筑大学经济与管理学院，研究方向为工程管理智能决策；*通信作者：孔梅（2001—）女，汉族，山东曲阜人，硕士在读，吉林建筑大学经济与管理学院，研究方向为建设工程管理。

Copyright © 2026 by authors and Viser Technology Pte. Ltd.