

## 人工智能新时代下教学改革的新思考

翟红林\*, 张晓昀, 曹晶晶

兰州大学化学化工学院, 兰州 730000

**摘要:** 随着ChatGPT-4的推出, 人工智能技术的发展与应用进入了一个崭新的时代, 也必然对现代科技人才的培养产生深远的影响。本文首先简述了人工智能技术及其在化学研究中的应用; 其次, 探讨了在人工智能新时代下化学专业学生创新能力培养的新要求; 最后, 介绍了我院《化学信息学》课程教学改革的探索与实践。

**关键字:** 人工智能; 化学信息学; 教学改革与实践

**中图分类号:** G64; O6

## Re-Thinking about Teaching Reform in the New Era of Artificial Intelligence

Honglin Zhai\*, Xiaoyun Zhang, Jingjing Cao

College of Chemistry & Chemical Engineering, Lanzhou University, Lanzhou 730000, China.

**Abstract:** With the advent of ChatGPT-4, the development and application of artificial intelligence (AI) technology have entered a new era, inevitably exerting a profound impact on the cultivation of modern scientific and technological talents. This paper begins by offering an overview of AI technology and its application in chemical research. It then discusses the updated criteria essential for fostering innovation ability of chemistry students in this new age of AI. Finally, the exploration and practice undertaken in the teaching reform for cheminformatics are presented.

**Key Words:** Artificial intelligence; Cheminformatics; Teaching reform and practice

自从2017年5月围棋系统AlphaGo机器人以3:0击败了当时的世界围棋冠军以来, 人工智能的理论、技术与应用进入了一个蓬勃发展的新阶段。特别是, ChatGPT-4的推出引发了人们极大的震撼。大数据、大算力、大模型所导致的智能涌现使得人工智能从传统的辨识模式走向了生成模式, 因而受到了高度的关注。有人为此欢呼, 有人感到恐慌。毋庸置疑的是, 人工智能的快速发展已经在不同层面、不同程度上影响着人们生活、学习及工作的思维和习惯, 不仅正在深刻地改变着科学研究的范式, 也必然地促使我们对现代科技人才的培养进行新的思考。

### 1 人工智能与机器学习

人工智能(Artificial Intelligence, AI)是指用于模拟、延伸和拓展人类智能的理论、方法与技术, 其本质特征在于, 基于已有的信息资源、采取复杂的数据挖掘算法建立可用于预测分析的数学模型。机器学习(Machine Learning, ML)则是实现人工智能最重要的策略之一, 其发展可大致分为三个层面(图1):

收稿: 2023-08-08; 录用: 2023-09-01; 网络发表: 2023-09-14

\*通讯作者, Email: zhaihl@lzu.edu.cn

基金资助: 兰州大学本科专业课程教学团队建设立项资助(202102012)

1) 基于有限样本数据、主要采用数理统计等方法建立预测模型。从数据的预处理、特征信息的提取到模型的建立等各个环节与步骤均由人工完成设计，机器仅“依计而行”，也被称为浅层机器学习。

2) 利用大量样本数据、以神经网络(Artificial Neural Network, ANN)为主要基础的复杂模型建立。机器能够通过自己的学习实现从原始数据中完全自动地筛选、提取所需的特征信息并完成与目标的映射关联，因而也被称为深度机器学习(简称深度学习，Deep Learning, DL)。

3) 采用海量的大数据样本、以深度学习为基础建立大型或超大型模型。大模型不仅可以更加精准地辨识、预测目标，而且还可以通过自主的学习去探索更大的信息空间而具备新的技能，即具有生成特性。其在诸多领域的应用取得了前所未有的突破，对人类生产与生活的众多方面带来变革性影响，也让人们再一次感受到了危机。



图1 人工智能发展的三个层次

## 2 人工智能对化学研究的变革

传统认知上，化学工作是在有试管、烧瓶、分析仪器等等的实验室中进行的，而化学研究则是“Chem-is-try”。然而，受益于量子化学与分子力学的发展，以理论计算与模拟为主的计算化学(Computational Chemistry)取得了长足的进步，不仅能够从分子、原子水平对实验结果加以解释、提供理论支撑，而且还可为更加高效地开展实验工作提供有益的指导。另一方面，各种数理统计的建模与分析方法也被大量地应用于化学信息的挖掘，以揭示化学变化的内在规律、助力化学实验研究。本质上，无论是理论计算与模拟还是统计建模与分析，都可归属于各种机器学习方法与技术在化学研究中的应用。不可否认的是，随着大数据的快速扩充、大算力的持续增强、人工智能大模型的不断涌现，基于实验和物理模型的化学研究与基于机器学习的人工智能正逐渐进入深度的融合。

例如在合成化学中，采用逆合成分析法进行合成路线的设计是最常用的策略。然而，即便是具有丰富经验的合成化学家也难以完全掌握并清晰记住数不胜数的化学反应。尽管借助化学文献数据库(如Reaxys)，化学家们只需通过分子结构、化学反应的输入便可筛选并设计出可能的合成路线，但要发现真正适合的、高效的合成反应新路线，仍然要依赖于化学家的知识与经验。计算机辅助合成路线设计(CASP)的研究由来已久<sup>[1]</sup>，并取得了可喜的进展<sup>[2]</sup>。2016年，韩国蔚山国家科学技术学院(UNIST)公布了一款名为Chematica的逆合成分析软件(现已更名为Synthia™)<sup>[3]</sup>；经过不断的改进与训练，该软件可用于设计复杂天然产物的全合成路线，并且已在实验室中成功地验证了其中的三种合成方法<sup>[4]</sup>。2018年，上海大学和德国明斯特大学的研究人员报道了一款设计分子合成路线的人工智能新工具，只是基于已经报道的单步反应即可自行学习化学转化规则，并进行快速、高效的逆合成分析<sup>[5]</sup>。2019年，麻省理工学院(MIT)研究人员研发了一种结合人工智能方法设计合成路线及机器人执行的自动化合成平台，能够为设定的目标分子提出多条合成路线及其反应条件等，并根据合成的步数及预测产量推荐最佳路径；同时，该平台拥有一个灵活的机器人手臂，可执行所有合成操

作流程、实现自动化合成，已成功用于15个化学小分子药物的合成路线设计和自动化合成<sup>[6]</sup>。同年，英国格拉斯哥大学的Leroy Cronin课题组采用Chemputer系统进行三种药物的合成，其产率与文献报道的人工合成相当<sup>[7]</sup>。2020年，利物浦大学Andrew I. Cooper课题组研发的自动化合成机器人在八天内进行了688次连续反应(图2)，发现了光解水产氢的光催化剂<sup>[8]</sup>。2023年，中山大学陈语谦团队依据反应转化的简易机理提出了一种基于图神经网络的Graph2Edits，用于逆合成预测<sup>[9]</sup>。当然，相对于合成化学的复杂性、艺术性而言，目前的人工智能还难以满足人们所有的需求。



图2 自动化合成机器人

尽管尚处于起步阶段，以生成式为主要特征的大模型构建与应用成为了人工智能发展的重要方向。化学的魅力之一在于不断的发现与创新，但以往的研究思路仍然局限于已知的化学事实，而这仅仅是化学信息空间中极其有限的一小部分。由于能够突破原有训练样本数据的局限而探索更大的信息空间，基于各种变分自动编码器(VAE)的生成式模型已经在分子设计与优化<sup>[10,11]</sup>、化学反应研究<sup>[12]</sup>、材料发现<sup>[13]</sup>等方面显示出了巨大的应用前景。

2021年5月，上海交通大学联合*Science*杂志发布了全球最前沿的125个科学问题(“125 questions: Exploration and discovery”)。其中第10个问题是，人工智能会重新定义化学的未来吗?(Will AI redefine the future of chemistry?)

该来的终究要来，谁也阻挡不了时代前进的脚步。

### 3 人工智能新时代下的人才培养

2007年，图灵奖得主Jim Gray在NRC-CSTB大会上提出了科学研究的四类范式<sup>[14]</sup>：实验科学是第一范式，在研究方法上以归纳为主，带有较多盲目性的观测和实验；理论科学是第二范式，偏重理论总结和理性概括，在研究方法上以演绎法为主；计算科学是第三范式，主要根据现有理论的模拟仿真计算，再进行少量的实验验证；数据密集型科学即第四范式，以大量数据为前提，运用机器学习等数据挖掘技术从大量已知数据中探索未知(图3)。

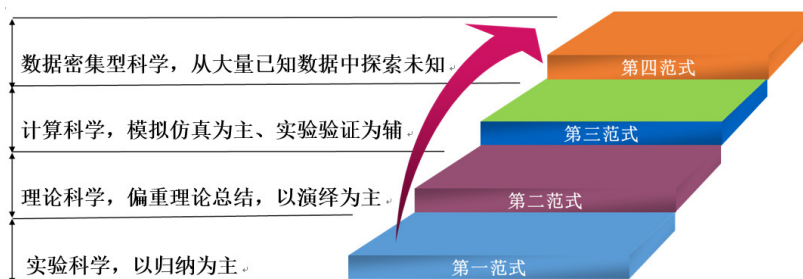


图3 科学研究四范式

显而易见，现代科技发展随着人工智能新时代的到来已经进入到第四范式，而化学研究则全面覆盖四个范式，形成了以实验、理论与计算三方面相互支持与促进的新特点。因此，对人才的培养提出了更新、更高的要求，主要体现在以下三个层面：

### (1) 注重信息素养、迎接AI挑战。

信息素养不仅仅是指能够进行计算机搜索、百度搜索、网上购物、智能支付等日常的简单操作，还应该在更高的水平上更深入地理解信息获取、转化、加工及应用的思维特点与技术实现，把握信息技术的优势与风险，并将其合理应用、有效地融入到生活与工作之中。这种思维习惯只能通过多维度、多层面的反复锤炼而养成。在这个大数据、大模型的新时代里，良好的信息素养已经成为现代专业人才极其重要的基本技能之一。因此，必须更加注重学生信息素养的提升，并贯穿于所有的教学活动之中。特别是面对AI大模型，要顺应时代的发展，引导学生迎接AI的挑战。

虽然AI技术的发展与应用让人瞠目结舌，但我们也应该清楚地认识到，真正的智能是以理解、分析为基础的感知与决策，而非仅仅是数据的驱动！人类决定性的主导地位难以撼动。当然，随着AI技术的不断进步，越来越多基础性、事物性甚至具有一定智力性的工作会由作为助手的AI取代完成；未来化学领域的AI应用可以将研究人员从繁琐的手工劳动中解放出来，使之有更多的时间和精力开展更高端的研发与创新。因此，AI技术的发展与应用不是人们躺平的理由，而是对人的能力提出了更高的要求。可以预见的是，AI加持下人所表现出来的能力才是新时代里最为重要的能力。

### (2) 更新教学思维、优化教学体系。

尽管我们都认同“授人以渔”、问题导向等教学理念，然而在实际的教学过程中却过多地关注了知识体系完整的传承，唯恐遗漏了某些重要的知识点而亏待了学生，从而导致了目前的教学体系大多仍以知识汇集为主。孰不知，知识是在源源不断地产生与更新的，学无止境；仅就掌握知识而言，人类难以超越经过训练的AI模型。能力的培养的确离不开知识点的支撑，但并非知识的丰富就直接等同于能力的高强。

教学讲授不同于学术报告，其关注点与方式方法各有不同。学术报告须体系完整、细节清晰；教学工作则必须要遵从人的认知规律、知识的传授应该是为学生思维的发展而服务的。我们以为，通过精心选取、有效组织少量的、必要的知识点为作支撑，着重于理论与方法、思考与判断的讲授及训练，引导学生自行延伸到其它更多相关的知识点。这样，不但可以减轻教与学的负担、缓解教学内容多与教学时数少的矛盾，更有利于激发、提高学生的学习热情和学习能力，拓展学科的交叉，提高教学效率。当然，这种新型教学体系的构建与优化是一个涉及多学科、多课程的系统工程，不仅要破除各学科课程间的人为界限，更需要加强思维发展的综合训练。这是教学改革中最重要、也是最大的改革方向，对教师的专业水平与教学能力提出了更高的要求。

### (3) 运用AI工具，助力教学改革。

习总书记强调，“中国高度重视人工智能对教育的深刻影响，积极推动人工智能和教育深度融合，促进教育变革创新”。国务院印发的《新一代人工智能发展规划》，明确利用智能技术加快推动人才培养模式、教学方法改革；教育部出台《高等学校人工智能创新行动计划》，并先后启动两批人工智能助推教师队伍建设的试点工作。7月8日，2023年世界人工智能大会教育论坛在世博中心举行，主题为“引领学习变革 智创教育未来”。可以肯定的是，人工智能技术的发展与应用将为教育教学改革提供巨大的空间及技术支撑；“人工智能+教育”在教、练、考、评、管各环节会不断碰撞出新的火花，将加速实现更加个性化、精准化、高效化的教育教学，为人才创新培养注入强劲动能。作为教师，我们必须积极顺应AI发展、充分学习并利用AI技术助力教学改革，以满足现代人才培养的需要。

在未来，AI不会取代教师职业，但很可能会取代不懂运用AI的教师。

#### 4 《化学信息学》课程教改的探索与实践

化学信息学作为化学专业人才培养的重要组成部分，一直受到我院的高度重视。随着入学大学生信息水平与能力的提高及信息技术类课程体系改革的深入，根据学院人才培养目标及化学信息学的发展，我们通过合并、撤换相关课程及内容，重新组建了《化学信息学》课程的教学体系<sup>[15]</sup>并编写出版了相应的本科生教材<sup>[16]</sup>。《化学信息学》内容繁多且涉及多学科的交叉，属化学专业信息素养类课程。经过几年的教学实践，针对《化学信息学》课程的特点，我们进行了多方探索并取得了良好的教学效果：

##### (1) 关注最新动态——追踪前沿。

化学信息学是信息学等相关技术与方法在化学领域中的应用，而新技术、新方法、新思维的不断出现，特别是各种AI大模型的持续推出，对化学学科的发展产生巨大的影响。作为《化学信息学》课程的教学，必须关注信息学方法及其在化学等领域中应用的研究进展及最新成果，并及时将其吸纳、融入到课堂教学之中，引领学生不仅了解信息技术的最新发展与应用，而且能够逐步地、更加深刻地认识与理解科技发展新时代下化学研究的新特点与新范式。如，随着ChatGPT-4的推出，适时地在课堂教学中介绍人工智能的新发展及其在化学研究中的新应用，引导学生紧跟时代发展，同时避免认知中的各种误区，正确理解AI大模型的运用及注意防范其带来的风险等。

##### (2) 提升信息素养——思维发展。

充分关注学生信息分析思维的建立、拓展及对信息处理的基本方法与工具的了解与掌握。在建模分析中，要让学生认识到从一阶数据到二阶数据可以描述更多、更全面的化学信息，同时也包含了更多的干扰。在数据处理方法上，从较为简单的一阶数据线性回归(多元、逐步、主成分、偏最小二乘等)入手，结合化学信息获取技术的进步，扩展到针对二阶数据发展起来的多道偏最小二乘、平行因子分析、多元曲线分辨-交替最小二乘等建模方法，再拓展到人工神经网络、深度机器学习。在分子模拟的教学中，使学生能够理解不同的基组、力场是基于不同的代表性样本集、通过不同的建模方法而获得，因而具有相应的适用体系及预测精度。通过教学使学生体会到建模分析的本质就是通过数据操作对化学信息进行变换以利于获得内在的关联或规律；人工智能、机器学习等各种模型的建立与应用乃至科学研究的基本思路与方法都类似于一个人的成长。所不同的是，人的成长是基于多层面、各角度、大跨度对社会与自然的综合感知与训练。

##### (3) 研究内容特点——分类施教。

化学信息学属多学科交叉，其特点之一是内容繁杂。在教学中，我们仔细地研究了不同内容具有的特点，采取不同的教学方式。如，对于化学信息学的起源与发展，采用在学生自行查阅相关资料的基础上，通过归纳总结引导学生理解科技的进步是需求与技术双轮推动的结果；在化学文献计算机检索部分，以学生实操教学为主，使学生在熟悉网站资源的同时，理解并掌握关键字检索与引文检索两大策略的特点与优势；在数据处理教学模块中，重点讲解化学信息的数据组织形式、处理与分析方法之间的关系及其实现，培养学生从实验、数据处理到结果分析的探究思维能力。此外，我们充分利用网络资源、线上教学平台、微信教学群等以实现不同教学内容的特性化教学。

##### (4) 加强实操训练——学以致用。

《化学信息学》课程的教学目标就是要在建立并强化学生专业信息素养的同时，进一步提升信息处理的能力，因而具有较强的工具类课程特征。通过实操训练，使学生易于理解相关的理论与方法。如在建模分析的课堂教学中，以Excel应用为主、MATLAB函数调用为辅加以实现，同时也鼓励学生使用他们熟悉或易学的SPSS、Python等工具。在ChemOffice及HyperChem软件的教学，除了分子结构、化学反应式等的绘制外，通过分子优化、单键旋转自由能变化、LUMO与HOMO能量、光谱模拟等计算，学生对于分子空间位阻、前线轨道理论、能级跃迁等有了更加形象、直观的认识，为相关课程的学习提供支持。与此同时，我们还注意将化学信息学的技术方法扩展到其它领域以拓宽学生的视野与思维、加强多学科知识的融合。如，主成分分析方法不仅可以用于化学数据，还可以

用于诸如大学排名、单位绩效考核、城市综合实力评价；因子分析方法还可应用于大气污染物的溯源；聚类与判别技术可用于新冠等诸多疾病的诊断等等。

对学生而言，掌握了有效的、好用的工具，就是为理想的实现插上了翅膀。

#### (5) 多种考核形式——全面客观。

根据不同教学内容的特点，采用不同的考核形式；在真实反映学生学习情况的同时，也引导、促进学生的学习。我们目前采用的考核形式主要有课堂提问(检查预习)、随堂测验(检查听课)、课后作业(复习、巩固与思考)、讨论课(拓展思维)、翻转课堂(检查课外小组学习成果)、课程小论文(考查信息处理综合能力)及期末考试等，实现过程性、阶段性学习的全面评价。另一方面，在对学生的考核过程中，教师也对教学工作进行不断地分析与反思并加以及时的调整；力争做到在因材施教的同时，对《化学信息学》课程体系进行不断的完善与优化。

## 5 结语

人工智能的发展无论是对人们的生活还是化学工作者的科研都必将产生深远的影响。在历次科技革命中，最终受益的都是敢于创新和拥抱变革的人，而那些忽视和拒绝接受改变的人，其结局往往会被时代抛弃。人工智能新时代的到来既是挑战、更是机遇：对于人才培养而言，信息素养的提升及思维能力的增强愈加重要；作为教师，要不断更新教学理念与思维、建立合理的课程体系，积极探索AI技术在教学活动中的应用，助力教学改革。

## 参 考 文 献

- [1] Corey, E. J.; Wipke, W. T. *Science* **1969**, *166* (3902), 178.
- [2] 丁邵珍, 江小琴, 孟超, 孙丽霞, 王正权, 杨弘宾, 沈国文, 夏宁. *中国科学: 化学*, **2023**, *53* (1), 66.
- [3] Szymkuc, S.; Gajewska, E. P.; Klucznik, T.; Molga, K.; Dittwald, P.; Startek, M.; Bajczyk, M.; Grzybowski, B. A. *Angew. Chem. Int. Ed.* **2016**, *55* (20), 5904.
- [4] Mikulak-Klucznik, B.; Golebiowska, P.; Bayly, A. A.; Popik, O.; Klucznik, T.; Szymkuc, S.; Gajewska, E. P.; Dittwald, P.; Staszewska-Krajewska, O.; Becker, W.; *et al.* *Nature* **2020**, *588* (7836), 83.
- [5] Segler, M. H. S.; Preuss, M.; Waller, M. P. *Nature* **2018**, *555* (7698), 604.
- [6] Coley, C. W.; Thomas, D. A., III; Lummiss, J. A. M.; Jaworski, J. N.; Breen, C. P.; Schultz, V.; Hart, T.; Fishman, J. S.; Rogers, L.; Gao, H.; *et al.* *Science* **2019**, *365* (6453), eaax1566.
- [7] Steiner, S.; Wolf, J.; Glatzel, S.; Andreou, A.; Granda, J. M.; Keenan, G.; Hinkley, T.; Aragon-Camarasa, G.; Kitson, P. J.; Angelone, D.; *et al.* *Science* **2019**, *363*, eaav2211.
- [8] Burger, B.; Maffettone, P. M.; Gusev, V. V.; Aitchison, C. M.; Bai, Y.; Wang, X. Y.; Li, X. B.; Alston, B.; Li, B. Y.; Clowes, R.; *et al.* *Nature* **2020**, *583* (7815), 237.
- [9] Zhong, W. H.; Yang, Z. D.; Chen, C. Y. C. *Nat. Commun.* **2023**, *14* (1), 3009.
- [10] Lim, J.; Ryu, S.; Kim, J. W.; Kim, W. Y. *J. Cheminform.* **2018**, *10* (1), 31.
- [11] Lee, M.; Min, K. *J. Chem. Inf. Model.* **2022**, *62* (12), 2943.
- [12] Kwon, Y.; Kim, S.; Choi, Y. S.; Kang, S. *J. Chem. Inf. Model.* **2022**, *62* (23), 5952.
- [13] Lyngby, P.; Thygesen, K. S. *NPJ Comput. Mater.* **2022**, *8* (1), 232.
- [14] Hey T.; Tansley S.; Tolle K. *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Microsoft Research: USA, 2009.
- [15] 翟红林, 张晓昀, 姚小军, 李书艳, 贾学庆, 曹晶晶. *大学化学*, **2019**, *34* (3), 11.
- [16] 翟红林. *化学信息学*. 北京: 化学工业出版社, 2019.