

■ 特稿

DOI:10.15998/j.cnki.issn1673-8012.2025.01.001

# 建设教育强国,未来学校何为



顾明远<sup>1</sup>,孔 苏<sup>2</sup>,郭 昶<sup>2</sup>,王 凯<sup>2</sup>

(1. 北京师范大学,北京 100875; 2. 杭州师范大学,杭州 311121)

**摘要:**为促进未来学校发展,培养更多具有创新意识和创新思维的人才,引导教师践行教育家精神,助力教育强国建设,顾明远先生深刻阐释了党的二十届三中全会和全国教育大会精神,认为教育要为培养高水平的科学技术人才做出贡献,要为全体人民的共同富裕打好基础。顾明远先生认为,教师的成长应当经历5项修炼,包括热爱教育、锤炼、学习、创新和收获。进一步推进素质教育,要注重培养学生的创新思维,要关注如何让学生从被动学习转变为主动学习。人工智能是信息技术发展的最新阶段,人工智能必然会改变我们的教育,改变教育生态和教育方式。利用信息技术,我们要处理好3组关系:一是目的与手段的关系;二是现代教育与传统教育的关系;三是虚拟世界与现实世界的关系。

**关键词:**教育强国;教育家精神;创新思维;未来教育;人工智能;个性化学习

[中图分类号]G640 [文献标志码]A [文章编号]16738012(2025)04000306

党的二十大报告指出,“教育、科技、人才是全面建设社会主义现代化国家的基础性、战略性支撑”。党的二十届三中全会提出,“要深化教育综合改革,深化科技体制改革,深化人才发展体制机制改革”。2024年9月,习近平总书记在全国教育大会上强调,要紧紧围绕立德树人根本任务,朝着建成教育强国战略目标扎实迈进。为更好助力教育强国建设,杭州师范大学基础教育发展中心未来学校研究所特邀请我国著名教育家顾明远先生莅临杭州师范大学附属凤鸣未来学校,对教育家精神、未来教育、未来学校等热点议题进行了深入阐述,为今后更好践行教育家精神、实施未来教育、办好未来

本文根据2024年9月28日顾明远先生在杭师大附属凤鸣未来学校的演讲与访谈整理,内容经顾明远先生审阅。

收稿日期:20241015

基金项目:浙江省哲学社会科学规划青年课题“基础教育优质均衡发展的浙江经验与先行示范研究”(24NDQN011YB)

作者简介:顾明远,男,江苏江阴人,北京师范大学资深教授,博士生导师,主要从事比较教育研究;

孔苏,男,安徽淮南人,杭州师范大学经亨颐教育学院讲师,教育学博士,主要从事教育基本理论研究;

郭昶,男,浙江宁波人,杭州师范大学附属凤鸣未来学校书记,正高级教师,主要从事学校管理研究;

王凯,男,湖北孝感人,杭州师范大学基础教育发展中心未来学校研究所所长,教授,教育学博士,主要从事课程与教学论研究。

引用格式:顾明远,孔苏,郭昶,等.建设教育强国,未来学校何为[J].重庆高教研究,2025,13(1):38.

Citation format:GU MINGyuan,KONG Su,GUO Chang,et al. On the future of schools in building a country strong in education [J]. Chongqing higher education research,2025,13(1):38.

学校提供了专业指导。

## 一、教育大计,教师为本,大力弘扬教育家精神<sup>①</sup>

### (一)深入学习领会党的二十届三中全会和全国教育大会精神

我们现在都在学习党的二十届三中全会精神和全国教育大会精神,这次全国教育大会的精神就是要建设教育强国。在2023年政治局会议上,习近平总书记强调,建设教育强国,是全面建成社会主义现代化强国的战略先导,是实现高水平科技自立自强的重要支撑,是促进全体人民共同富裕的有效途径,是以中国式现代化全面推进中华民族伟大复兴的基础工程。

科学技术是第一生产力,人才是关键,基础在教育。战略人才是我国高水平科技自立自强的有力支撑。建设教育强国是实现全体人民共同富裕的有效途径。因此,建设教育强国,教育要为培养高水平的科学技术人才做出贡献,要为全体人民的共同富裕打好基础。建设教育强国是通过中国式现代化来实现中华民族伟大复兴的中国梦。

建设教育强国,一是要优先发展教育;二是要培养高水平的科学技术人才即拔尖创新人才,同时培养实现共同富裕不可或缺的普通劳动者、大国工匠和现代农民。

在今天的全国教育大会上,习近平总书记强调,“我们要建成的教育强国,是中国特色社会主义教育强国,应当具有强大的思政引领力、人才竞争力、科技支撑力、民生保障力、社会协同力、国际影响力,为以中国式现代化全面推进强国建设、民族复兴伟业提供有力支撑”<sup>[1]</sup>。这为我们建设教育强国制定了标准,即怎么样做才能实现教育强国。

### (二)弘扬教育家精神,教师要经历五项修炼

教育大计,教师为本。教育战线的基石在学校、在基层、在教师。党中央高度重视教师队伍建设,习近平总书记在2018年全国教育大会上指出,坚持把教师队伍建设作为基础工作。2014年9月9日,习近平总书记在同北京师范大学师生代表座谈时,提出有理想信念、有道德情操、有扎实学识、有仁爱之心的“四有”好老师标准。2016年9月9日,习近平总书记在与北京市八一学校师生座谈时,鼓励教师做学生锤炼品格的引路人、学习知识的引路人、创新思维的引路人、奉献祖国的引路人。2021年4月19日,习近平总书记在考察清华大学时提出教师要做“大先生”。2023年9月9日,习近平总书记在致信全国优秀教师代表时提出中国特有的教育家精神——心有大我、至诚报国的理想信念,言为士则、行为世范的道德情操,启智润心、因材施教的育人智慧,勤学笃行、求是创新的躬耕态度,乐教爱生、甘于奉献的仁爱之心,胸怀天下、以文化人的弘道追求。

教师的成长,要经历5项修炼。

第一项修炼是热爱教育。这是对教师最基本的要求,也是教育发展的基础。教育工作为什么重要?因为教育工作首先面对的是青少年儿童。我们要关心每一个孩子,因为这关系到他的一生,关系到他的家庭幸福。说大一点,它关系到我们民族的未来和国家的兴衰。我们要实现中国式现代化、实现教育现代化,首先要培养现代人。现代人具有民主、自由、高素质等特征。教师对教育要有这样的认识,对自身职业也要有认识。教师直接面对青少年儿童,不仅要为他们传授知识,更重要的是培养他们成人。就像陶行知先生讲的,“千教万教教人求真,千学万学学做真人”。怎么成为真人?要有高尚的品德,要胸怀祖国,要有真实本领、勇于担当。我们做工业、做农业都需要其他手段,教师职业不是靠其他的手段,而是要直接针对儿童的心灵,这是所有其他职业都取代不了的。

<sup>①</sup> 本节是顾明远先生关于学习党的二十届三中全会和全国教育大会精神的主旨演讲内容,标题为新增,文字有调整。

第二项修炼是锤炼。我们这些老教师都是锤炼出来的,都是在教育里面犯错误、从不成熟走向成熟。新教师不是一毕业走上讲台就能上好一堂课的。就像于漪老师讲的,上好一堂课,备课是一辈子的事。要不断磨炼,好课都是磨出来的。上完课以后,哪些内容讲得好、哪些地方还有不足,要做好反思。经过锤炼,首先成为合格的教师,再逐步成为优秀教师。

第三项修炼是学习。我们遇到问题怎么办,就是要学习。向书本学习,学习一些教育理论,还要学习教育的历史。我们五千多年的教育是怎么演化的,国际上有什么经验,历史上有什么经验,包括对一些教育问题的争论,都要学习。我想,最主要的还是要提高教师的专业水平。专业水平包括自身的修养,中国人经常讲修养身心,一个教师有修养就有教育智慧,所以要学习。不光遇到问题要学习,日常也要学习。我经常倡导,文科教师要多学习科普知识,当代科学技术日新月异,不了解一点科技知识,对学生的教育难以取得良好的效果。同样,理科教师也要学习一些文艺知识,提升自我的文化修养。有了修养就有悟性,智慧往往与悟性相关。

第四项修炼是创新。教育事业是创新的事业,我们天天讲教育改革,每一天可能都会遇到新问题。一个班级昨天挺好,今天就出现了一些问题;有的学生昨天心情挺好,今天心情就不太好。这个时候教师就要细心,所以我曾讲“教书育人在细微处”。要创新,每个教师就要有自己的教育风格。教师本身的素质是不同的,有的教师逻辑性比较强,有的教师形象思维比较好,要根据自己的情况创造自己的风格。每个教育家、优秀教师,都有自己的风格。

第五项修炼是收获。所谓收获,就是看到儿童的成长,看到他一天一天成人,这就是教育的成功,也是教师的成功,教师就有收获的喜悦。这种收获也要修炼,因为这种收获是不求回报的,跟父母要求孩子不一样。我们的收获是孩子的成长,是对社会服务,对国家、对民族、对未来的爱的体现。我们的优秀教师要更加努力培养下一代优秀教师,使我们的“优秀”能够一代一代传承下去。一个优秀教师的成长,可能需要经过这5个方面的修炼,慢慢成为教育家。

### (三)以教育家精神赋能校长和教师工作

作为一个校长,弘扬教育家精神就是要做好本职工作,办好每一所学校,教好每一个学生。校长要走进课堂,关心教师,关心学生。弘扬教育家精神,不要停留在口头上,要践行在自己的工作中。

教育家,其实成长在课堂上。我曾经说,教师上好每一堂课是很重要的,因为教学工作是学校工作的核心,上课是主要的教学工作,学生在学校的主要活动都发生在课堂上<sup>[2]</sup>。我们讲教育改革要创新,如何创新?我认为,要转变我们的观念。当前的教育,已经不是原来那样教师抱着课本讲好现成的知识就行了。现代教育有了科学技术特别是人工智能的辅助,正在改变我们的教育教学方式,也改变着师生关系。所以当前我们要改革创新,有一条很重要,就是让学生从被动学习转变为主动学习。现在有些孩子学习不是很积极,有些孩子对学习没有兴趣,怎么让学生主动学习?怎么变被动学习为主动学习?特别是现在有些家长强迫孩子上各种补习班,学生负担过重,学生被动学习的问题还没有根本转变。我们要让学生自主地学习、有兴趣地学习。所以我经常讲,没有兴趣就没有学习。

教育家精神是新时代对教师提出的更高要求。我们的教育已经进入一个新的时代,已经普及了义务教育、高中教育和高等教育,我们的教育要从数量的发展转变到质量的提高上来。习近平总书记一直强调,要建设一支高素质专业化的教师队伍,这支教师队伍就是具有教育家精神的教师队伍。这支队伍的教师不可能人人都是教育家,但是可以以教育家精神来锤炼自己,使人人都争取成为教育家。

教育家精神有6条明确标准,用简单的4个字概括就是“敬业爱生”。敬业,就是对教育事业有高度的认识,勤勤恳恳,一辈子为教育事业做贡献;爱生,就是要关爱每一个学生,不放弃每一个学生。

## 二、改革创新,引领未来,促进教育高质量发展<sup>①</sup>

### (一)进一步推进素质教育的关键是改革创新

素质教育正式提出已经很多年了。要真正理解素质教育的精神实质,最重要的是解放思想,从传统教育观念的束缚中解放出来<sup>[3]</sup>。素质教育就是要五育并举,把学校、家庭、社会统筹起来培养学生,提高学生的思想品德和身体素质,同时,培养学生具有扎实的学识。不仅仅是让学生获得一些知识,更要培养学生的创新思维,使学生具有高尚的思想品德和良好的综合素养。

现在我们提倡“双减”,很多家长不理解,认为减轻负担以后,学生的学习质量会下降。实际上相反,我们减轻负担的目的就是提高学生的学习质量。学习负担减轻了,学生学习有兴趣了,他的学习效率就会提高。学习负担减轻了,学生就有时间走向大自然,走向社会,参加他所喜爱的活动。前一个阶段,人人学奥数,这对一些数学没有兴趣和天赋的孩子,是一种不公平的教育。奥数并不是坏事,但它只适合少数对数学有兴趣或有天赋的学生。让所有学生都学奥数,这是对一部分学生最大的不公平。减轻负担以后,不再用一个标准评价学生,学生就有时间参加自己喜爱的活动,他喜欢数学就学奥数,喜欢艺术可以去学舞蹈,等等,这样才能真正培养出具有个性的人才。我们现在提出要培养拔尖创新人才,怎么培养?只有在普及、普遍的基础上,发现孩子对某种学科有兴趣、有天赋,教师及时引导、因材施教,拔尖创新人才才能脱颖而出。过去一堂课就是一个标准,这不可能培养出拔尖创新人才。教师要细心观察孩子的喜好和天赋,决不能拔苗助长。“双减”就是让所有孩子发挥自己的才能,从中发现一些特殊人才。

我们的教育改革,往往都走到应试教育的轨道上了,这样的改革不会有多少成效。如果不走出应试教育的阴影,就很难培养出拔尖创新人才。考试只是手段不是目的,平常的考试是对学生学习情况、教师教学情况的检查,中考、高考作为选拔性考试,是基于日常的学习成绩、考试情况,不能专门为了中考、高考而教学。我们在应试中只有智育,其他都没有关注到。现在是创新的时代,改革创新要培养学生的创新思维。不创新,我们的国家就无法立足于世界民族之林。西方国家制裁我们,妄图使我们的科技不要发展,不要超过他们。他们封锁我们,我们怎么办?党中央提出高水平科技自立自强,这对基础教育意味着什么?就是要培养学生的创新思维和创新能力。过去的应试教育压抑了学生的创新思维。本来孩子有创新的天性,小学一年级、二年级的孩子喜欢问很多问题,对世界充满了好奇,但经过我们的教育,他的好奇心不是增长了,而是减少了。从早到晚看书本、做作业,压抑了好奇心和创新思维。所以,我们还是要加强素质教育,重视德智体美劳五育并举。

### (二)基于人工智能的未来教育发展方向

什么是未来教育?有的学者从科技角度讲,就是适应科学技术发展的要求,这当然是对的。任何一种生产力,都会影响教育。孔子讲学的时候,没有教材和课本,字都是刻在竹简和木简上。后来出现了纸张,可以更好地传播。再后来发明了印刷术,书都是印刷的,就更加容易传播了。工业革命以后才有了义务教育。工业革命以前,教育没有普及,大多数人都是文盲。当前,出现了信息技术和人工智能,人工智能是信息技术发展的新阶段,它必然改变我们的教育,改变教育生态和教育方式。

教育都是面向未来的,都是为未来社会培养人才,所以未来教育不是一个新名词。当然,现在说的未来教育有一些新含义,比如科学技术特别是人工智能、大数据影响学校发展。但学校未来发展的根本还是立德树人,我们的未来教育还是要着眼于培养实现中华民族伟大复兴的一代新人,现在的学生就是促进中华民族伟大复兴的新生力量。未来教育不能只看到技术,其逻辑指向还是在面对未来

<sup>①</sup> 本节是顾明远先生接受杭州师范大学未来学校研究所专访的内容,标题为新增,文字有调整。

社会。未来社会存在很多不确定性,但有一点是确定的,就是要创新。未来教育的核心就是创新。

人工智能要赋能教育,促进教育发展,需要分层次。第一个层次是教育管理,包括教育的区域管理。第二个层次是教学工作,最主要的还是课堂。人工智能可以促进教育公平,但做不好也会拉大教育差距。人工智能只是手段,要善于利用。我们对人工智能的认识还不够,教育领域对人工智能的利用也远远落后于其他领域。教育不仅是技术问题,更是培养人、改变心灵、塑造灵魂的问题,这是机器不可替代的。教育部建立了国家智慧教育公共服务平台,很多优质教育资源可以共享,以促进教育公平。但在一些地区和学校,信息技术还停留在制作课件、查找资料层面。

充分认识“人工智能+教育”的育人功能是当前教育工作者遇到的重大挑战<sup>[4]</sup>。信息技术、人工智能对教育的最大影响是可以实现因材施教和个性化培养。过去一个班有几十个孩子,很难做到个性化培养。但现在有人工智能和大数据支持,我们就可以了解每一个孩子的学习状况,他在课堂上干了什么、看了什么、学习的情况怎样、情绪怎样。当教师了解到这些信息的时候,就能够做到因材施教,为每个孩子制定个性化的学习方案,包括向学生推荐书籍和学习内容等,促进学生的个性发展。我们要充分利用人工智能、大数据的优势,还要关注互联网的互联性、开放性和虚拟性。过去与学生沟通是面对面的,现在利用互联网,我们在手机上就可以与学生沟通。所以要充分利用信息技术,实现更多的师生互动。

利用信息技术要处理好3组关系:一是目的与手段的关系,也就是人文主义与技术主义的关系。人工智能是手段,立德树人是目的,这个不能忘,也不会变。不是机器代替人,而是让机器更好地促进人的发展。过去讲,有了人工智能就不用黑板了。我提出,还是要留一块黑板,老师在黑板上画一画、写一写,就有人情味。这和PPT课件不一样。二是现代教育与传统教育的关系。现代教育可以利用信息技术,但传统教育的优势不能丢。有些课本内容,还需要老师讲一讲。过去认为有了PPT,老师就不用讲了。但我认为,老师还是要把基本的概念好好讲一讲。什么是一堂好课?简单地说,第一,能不能启发学生的思维。是把现成的知识告诉学生,还是告诉学生这个知识是怎么来的,这个概念是怎么形成的,这个学科将来的发展怎么样,还有哪些是我们不清楚的,这很关键。第二,能不能关照到每一个孩子。我经常去听课,发现有些教师只喜欢聪明的孩子,提问总找聪明的孩子回答。有的孩子比较内向、木讷,不积极举手发言,那么,这样的孩子如何成长?三是要处理好虚拟世界与现实世界的关系。现代教育可以利用技术手段呈现很多虚拟的内容,但现实世界更复杂,要让学生走向自然、走进社会,去实际体验真实的世界。

### (三)加强中华优秀传统文化教育

中国的教育是中国共产党领导的社会主义教育,我们要培养中国特色社会主义道路自信、理论自信、制度自信和文化自信的接班人,就必须重视中华优秀传统文化教育。中华优秀传统文化是我们民族的根,也是我们教育的根和魂。一个民族如果失去了它的民族精神,这个民族也就不复存在了。当前我们在吸取西方优秀教育理念、教育方法的同时,一定不能忘记中华优秀传统文化的精神。中华优秀传统文化的精神流淌在每个中国人的血液中,我们要从教育方面把它加强。我曾提出,开展中华优秀传统文化教育应当注意4个问题。首先,我们要有所选择。其次,开展中华优秀传统文化教育,要以课堂教学为主。再次,课外和校外活动也是进行中华优秀传统文化教育的有效途径。最后,进行中华优秀传统文化教育要重视它的基本精神,避免形式主义<sup>[5]</sup>。目前,我们正在以中国式现代化全面推进中华民族伟大复兴,教育对中华优秀传统文化的传承和发扬显得尤为重要,也更加迫切。

### (四)未来教育、终身教育与学习型社会密切相关

有人讲,未来教育就是终身教育。终身教育的提出,尽管只有几十年时间,但它是实现可持续发展的一个很重要的手段。1993年,中共中央、国务院发布《中国教育改革和发展纲要》,第一次在国家

文件中提出要建立终身教育体系。社会要持续发展,就需要人人学习、人人适应时代的变化。所以,未来教育也是终身的教育。1972年,联合国教科文组织国际教育发展委员会的报告《学会生存——教育世界的今天和明天》指出,未来的教育必须成为一个协调的整体,在这个整体内,社会的一切部门都从结构上统一起来。这个教育运动是不可抗拒的和不可逆转的。我曾提出,在未来社会里每个成员都是学习者,每一个社会组织都是学习型组织<sup>[6]</sup>。从学校来说,未来教育主要是培养学生的未来意识,未来意识就是创新意识、可持续发展意识。同时,要培养他们终身学习的意识,而不是大学毕业了,在学校学完了,学习就结束了。作为学校,还应当帮助社区居民终身学习,适应社会发展变化的需要。终身教育是推动社会成为学习型社会的关键力量。

### 参考文献:

- [1] 习近平在全国教育大会上强调 紧紧围绕立德树人根本任务 朝着建成教育强国战略目标扎实迈进[EB/OL]. (20240910)[20240929]. [http://www.moe.gov.cn/jyb\\_xwfb/s6052/moe\\_838/202409/t20240910\\_1150246.html](http://www.moe.gov.cn/jyb_xwfb/s6052/moe_838/202409/t20240910_1150246.html).
- [2] 顾明远,滕珺,韩思阳.上好每一堂课 教好每一个学生:顾明远先生教师教育思想专访[J].教师教育学报,2021(2):48.
- [3] 顾明远.教育公平与素质教育[J].教育发展研究,2002(1):1011.
- [4] 顾明远.“人工智能+教育”将使教育发生革命性的变革[J].教育家,2019(40):1.
- [5] 顾明远.文化是一个民族的根与魂:谈谈中华优秀传统文化教育[J].人民教育,2017(23):4549.
- [6] 顾明远.形成全民学习、终身学习的学习型社会[J].求是,2003(4):42.

(责任编辑:蔡宗模 校对:杨慷慨)

## On the Future of Schools in Building a Country Strong in Education

GU MINGYuan<sup>1</sup>, KONG Su<sup>2</sup>, GUO Chang<sup>3</sup>, WANG Kai<sup>4</sup>

(1. Beijing Normal University, Beijing 100875, China; 2. Hangzhou Normal University, Hangzhou 311121, China)

**Abstract:** In order to promote the development of future schools, cultivate more talents with innovative consciousness and thinking, guide teachers to practice the spirit of educators, and assist in the construction of a country strong in education, Mr. Gu Mingyuan deeply explained the spirit of studying the Third Plenary Session of the 20th Central Committee of the Communist Party of China and the National Education Congress. He believed that education should contribute to the cultivation of high-level scientific and technological talents, lay a solid foundation for the common prosperity of all people. Mr. Gu Mingyuan believes that the growth of teachers requires five practices, including love for education, refinement, learning, innovation, and achievement. Contemporary educational reform and innovation should focus on cultivating students' innovative thinking and how to transform them from passive learning to active learning. Artificial intelligence is the latest stage of information technology development, and it will inevitably change our education, our educational ecology, and our educational methods. Using information technology, we need to handle three sets of relationships: firstly, the relationship between purpose and means; secondly the relationship between modern education and traditional education; thirdly the relationship between the virtual world and the real world.

**Key words:** a country strong in education; the spirit of educators; innovative thinking; future education; artificial intelligence; personalized learning

教育数智化

DOI:10.15998/j.cnki.issn1673-8012.2025.01.002

# 人工智能赋能高校科学研究范式创新： 价值、风险与进路



赵晓伟,王小雨,王艺蓉,沈书生

(南京师范大学 教育科学学院,南京 210097)

**摘要:**高校是科学研究的主阵地,是基础研究与重大科技突破的主力军,肩负着加快形成新质生产力的重要使命。在新一轮科技革命浪潮下,人工智能技术在科学研究中的广泛应用与深度融合,不仅重塑了科学研究的基本过程,更引领着科学研究范式向智能科学研究范式转型。从逻辑实证主义视角出发,探讨人工智能赋能高校科学研究范式创新的多重价值,包括增强科学研究中的数据采集、加速科学假设的生成与验证、实现实验模拟与自动化,以及激发科学洞察与创意涌现,有助于理解AI赋能科学研究范式的价值和深远影响。基于此,审视当前科研范式转型过程中遭遇的诸多挑战与风险,包括人的边缘化风险(“人在旁路”现象)、数据偏见与误导性主张的双重困境(“数据陷阱”问题)、算法透明度的缺失(“算法黑箱”问题)以及由此引发的信任危机等,高校在科学研究范式转型中应紧密依托新一轮科技变革,植根中国科学研究实践土壤,立足已有范式寻求创变,探索一条符合我国国情且具有鲜明中国特色的高校科学研究道路。具体来讲,应加强跨学科合作与协同创新、规范数据使用标准、推动可解释性AI的发展,建立健全科研伦理审查机制,为科研范式创新提供坚实的理论支撑与行动指南。

**关键词:**科学研究范式;智能科学研究范式;人工智能;数据陷阱;信任危机

[中图分类号]G644 [文献标志码]A [文章编号]16738012(2025)010009-12

修回日期:20240815

**基金项目:**国家社会科学基金重大项目“新一代人工智能对教育的影响研究”(VGA230012);认知智能国家重点实验室智能教育开放课题“新一代人工智能促进教育管理变革与政策创新”(iED2024M002)

**作者简介:**赵晓伟,女,山东莒县人,南京师范大学教育科学学院讲师,博士,主要从事人机协同教育;

王小雨,女,江苏兴化人,南京师范大学教育科学学院硕士生,主要从事信息化教学设计;

王艺蓉,女,浙江杭州人,南京师范大学教育科学学院硕士生,主要从事教育技术研究。

**通信作者:**沈书生,男,江苏海安人,南京师范大学教育科学学院教授,博士生导师,主要从事教育技术理论研究。

**引用格式:**赵晓伟,王小雨,王艺蓉,等.人工智能赋能高校科学研究范式创新:价值、风险与进路[J].重庆高教研究,2025,13(1):920.

**Citation format:** ZHAO Xiaowei, WANG Xiaoyu, WANG Yirong, et al. Empowering innovation in scientific research paradigms at higher education institutions with artificial intelligence: value, risks and approaches[J]. Chongqing higher education research, 2025, 13(1): 920.

电子显微镜的发明使人类首次发现细胞内部的超微结构,大型强子对撞机的制造使希格斯玻色子的存在得以证实。不难发现,历次重大科学发现都离不开新工具或新技术的诞生。新一轮科技革命加速了科学研究的进程,从根本上改变了科学研究的方式,推动科学研究的范式转型。在数据密集型第四科学研究范式下,研究者依托庞大的数据集进行数据建模与分析,探索数据中的潜在规律。以生成式 AI 为代表的新一代 AI 技术,不仅能够高效完成大规模的数据计算,更能协助研究者发现潜在的科学问题,探寻新的研究方向,为科学研究提供新方法,推动科学研究向第五范式——智能科学研究范式(AI for Science, AI4S)转型。

AI4S 智能科学研究范式代表着 AI 数智技术与科学研究的深度融合和双向赋能<sup>[1]</sup>,加速科学发现与研究的进程,实现科学研究过程的自动化和内容生产的智能化。与此同时,科学研究的发展推动着智能算法模型的持续优化与迭代。高校是科学研究的主阵地,高校研究人员是基础研究与重大科技突破的主力军。高校肩负着加快形成新质生产力、实现高水平科技自立自强、推动科学研究高质量发展的重要使命,应响应国家科技战略需要推进科学研究创新,以提升国家的原始科技创新能力。在新科技革命背景下,基础研究转化周期明显缩短,国际科技竞争形势激烈,人工智能驱动科学研究不仅跃升为政策制定者与战略规划者无法回避的核心议题,更是推动社会进步的关键。高校应以数智技术作为科学研究范式转型的主要引擎,推动人工智能与科学研究双向赋能,发挥人工智能在科学研究及其范式创新上的强大驱动力<sup>[2]</sup>。本研究梳理了 AI 赋能高校科学研究范式创新的价值与风险挑战,并针对性地提出应对策略,以期为高校科学研究范式的变革提供参考。

## 一、AI 赋能高校科学研究范式创新的价值意蕴

库恩指出,科学革命就是范式的更替,范式的确立与科学共同体的行为准则及其待解的谜题紧密相连,涉及某些实际科学实践的公认范例,为特定科学研究提供参照<sup>[3]</sup>。科学研究范式不仅是科学发现与科技创新的基石,更是科学共同体在特定时期进行科学研究的方式,契合科技创新的内在规律<sup>[4]</sup>。科学研究实质上是追寻“自然之谜”答案的过程,逻辑实证主义者认为这一过程遵循既定的科学探究模式。亨普尔(Hempel)将这一过程划分为 4 个阶段:观察记录事实、分析归类事实、归纳推导结论和检验结论<sup>[5]</sup>。科学发现被视为从提出假说至被接受的过程,Laudan 认为,此中间过程不仅是归纳,更需要在猜想或假设后进行求证与完善<sup>[6]</sup>。整体来看,科学发现大致包括观察提问、假说建立、实验验证、理论构建并持续修正的过程。

科学研究范式的演进经历了多个阶段,从实验描述的经验范式,到模型归纳的理论范式、仿真模拟的计算范式,再到数据密集型分析的数据范式,直到如今的智能科学研究范式(AI4S)。在新科学研究范式下,AI 增强了研究者的数据获取方式与计算能力,深刻改变了假设建立的逻辑,通过自动化或模拟实验过程,加速科学洞见的形成,展现了其强大的潜力。同时,AI 通过回溯、预见与迭代等方式,不断优化科学研究的过程,推动科学生产力的提升,基本重塑了科学研究的一般过程(如图 1)。

### (一) 助益科学研究中的数据采集

在科学研究领域,数据的采集转换与深度分析是构建科学见解和理论的基石,AI4S 背景下高校科学研究在数据采集与管理的各个环节已取得突破性进展。在数据发现与选择方面,深度自动编码器和其他先进的深度神经网络已展现出从庞大且结构复杂的研究数据中提取非线性特征、筛选高价值研究训练数据的能力。这些智能技术能够根据研究需求,自动识别和剔除异常值,获取更为纯净的数据,有效应对存储与传输大规模研究数据的技术挑战<sup>[7]</sup>。在数据标注方面,大型标签数据集对于模型的训练至关重要,随着标注策略的不断革新,自监督学习等先进方法已能够实现标签的自动标记,极大地减轻了研究者在数据标注上的负担,能够以更少的数据标签完成数据集的全面标注,为模

型训练提供有力支持。在数据生成方面,随着研究数据集质量与规模的不断提升,用于数据增强和模型训练的算法也持续迭代,不仅能够扩充数据集规模,还能集成不同类型的数据集,生成跨学科知识领域的的数据,从而大幅提高数据集的综合性<sup>[8]</sup>。此外,AI的发展还优化了数据细化处理的能力,提升数据质量的精度,为科学发现与洞见提供可靠的数据支撑。



图1 AI赋能高校科学研究范式的基本逻辑

芝加哥大学数据科学研究所 Willett 提出,科学家可以借助 AI 分析复杂的科学数据,如借助机器学习与深度学习技术,从韦伯望远镜捕获的宇宙图片中挖掘出潜在的高价值数据,从而揭示宇宙的奥秘<sup>[9]</sup>。在物理科学实验中,AI 能够辅助科学家对海量实验数据进行筛选、采集与存储。如在质子碰撞试验中,利用基于深度学习的自动编码器对碰撞产生的数据进行过滤性采集,能够有效应对数据存储的极限挑战<sup>[10]</sup>。算法模型的训练往往需要专业人员对庞大的数据集进行标注,而垂直领域模型的通用性和推广性往往受到限制,因此数据标注工作尤为繁重。伦敦大学学院和 Moorfields 眼科医院提出了一种名为 RETFound 的视网膜图像基础模型,该模型利用自监督学习,在超过 160 万张未标注的视网膜图像上进行训练,并在眼部疾病诊断、预后分析以及系统性疾病预测等任务中表现出卓越性能<sup>[11]</sup>。该成果不仅减轻了数据标注的负担,更展示了 AI 在科学研究领域的巨大潜力和广阔前景。

## (二)加速科学发现中的假设生成

在科学发现的过程中,提出可供检验的研究假设是重要开端。传统科学研究范式下这一过程往往费时费力,深受研究者的专业知识、探索能力以及个人偏见的影响。AI4S 背景下,数智技术为假设生成方式带来革命性改变,提升了生成效率,推动科学见解的整合,加速推进隐式和未解的科学发现。假设生成方式主要包括两类:

一是基于文献的知识发现(literature-based discovery, LBD)。该方法运用文本挖掘、自然语言处理及机器学习等技术为模型提供测试数据集,对海量文献数据进行深度处理,经训练的模型能够自动从文献中抽取知识概念及隐含关联,助力研究者高效提出研究假设。目前,LBD 在生物医学、材料科学等领域已取得显著成效,预计下一代模型将具备处理图表、代码等多模态研究数据能力。例如,劳伦斯伯克利国家实验室 Tshitoyan 团队采用自然语言处理的无监督学习 AI 技术,分析材料科学领域的论文摘要,发现训练后的系统具备“化学直觉”,能够预测具有特定属性的新材料,并提出可能材料的假设与预测<sup>[12]</sup>。芝加哥大学社会学家 Sourati 等人进一步扩展了该方法,并将论文作者间的关联纳入考虑,训练了兼顾论文摘要与作者的 LBD 系统,发现该系统预测材料的能力达到 Tshitoyan 团队的

两倍<sup>[13]</sup>。

二是智能辅助的假设空间聚焦<sup>[8]</sup>。主要包括 3 个过程:首先,假设对象的快速筛选。科学研究中一个科学问题往往包含多个可能的假设对象或符号,构成假设空间时并非所有假设都值得深入探索, AI 将辅助科研人员快速筛选对象或符号,聚合最有可能产生价值的假设范围。例如,利用自监督学习的 AI 预测器,可以根据已有知识和数据,预测哪些对象或符号更有可能产生有意义的假设。其次,提升假设的质量和生成效率。一旦确定要探索的假设空间,就需要生成具体的假设, AI 可以通过优化算法和强化学习技术,助力生成高质量的假设。具体而言,可训练 AI 代理(如强化学习模型),在假设空间中寻找高质量的候选假设,该代理将根据需要采用“奖励策略”的方法,不断尝试和改进,直到找到满足要求的假设。最后,实验前的假设评估与预测。在确定待探索的假设后,还需要对其加以评估和预测,以确定哪些假设最值得进行实验验证。该过程同样可以借助 AI 的支持,通过整合先前的知识和数据,学习研究假设的贝叶斯后验分布(即假设在给定数据下的概率分布),可使用预测模型对假设进行实验前的对比和评估,选择最有可能成立的假设进行实验。该方法可以帮助研究者更准确地预测实验结果,减少实验浪费,提高科学研究效率。

### (三)实现科学实验模拟与自动化

实验作为检验科学假设的关键,是推动科学发现的核心环节。传统实验面临成本高昂、效率低下、精度不足、难以实施等诸多困境,限制了科学研究的步伐。AI4S 背景下科学实验迎来模拟化与自动化的革命,不仅降低实验成本、提高效率,更以空前的规模自动化开展,拓展科学实验边界的同时,更在深层次上重塑科学实验的本质。科学实验模拟化即利用计算机模拟科学实验的全过程,得益于高性能计算机与深度学习技术的进步,大模型、大数据与大算力的融合,科学实验朝着计算机模拟的方向飞速发展,为复杂实验的实施开辟了新途径,目前已在多领域取得显著成效。如在气象科学领域,DeepMind、华为和 Nvidia 等科技公司推出了 AI 天气预测模型,不仅提高预报的速度和准确度,更降低了预测成本。Nvidia 公司更是计划通过“Earth-2”项目开展更大区域的气候预测,以展现 AI 在气象科学中的巨大潜力<sup>[14]</sup>。在化学领域,麻省理工学院研究人员通过机器学习模型成果模拟了化学反应中的过渡态这一短暂过程,加速了化学反应物和催化剂的研究进程<sup>[15]</sup>。

科学实验自动化的实现则依托大语言模型、人工智能、机器人等先进技术的支持。在该模式下,大型科学研究平台或自动实验系统能够独立设计实验并执行,缓解人类实验员的压力。从实验前的方案制定、方法选择和环境条件配置,到实验中的数据记录、参数监控与调整,再到实验后的数据分析与结论生成, AI 均能胜任。这一变革极大地解放了科学研究的生产力,科学家们能够以前所未有的速度完成实验。例如,密歇根大学研发的 BacterAI 平台,能够实现细菌实验的高度自动化,每日可完成上万次实验,全程无需人工介入<sup>[16]</sup>。中国科学技术大学研发的机器人化学家“小来”,由“化学大脑”、机器人实验员和化学工作站组成,其中“化学大脑”能阅读文献、设计实验、优化方案,机器人实验员与工作站协同执行实验,在短短 5 周内完成传统方法需耗时 1 400 年的化学实验<sup>[17]</sup>。自动化实验的另一优势在于高效执行重复性实验,科学界普遍存在对重复他人工作的动力不足的现象,自动化实验机器人能够对先前研究成果进行重复性实验,节省研究者的精力,不仅验证已有科学主张,还详细记录实验步骤,为后续学术分析提供宝贵数据,推动科学研究的可持续发展。例如,曼彻斯特大学 Roper 教授团队借助 Eve 机器人成功复制多项生物医学研究成果,为科学发展注入新活力<sup>[18]</sup>。

### (四)激发科学预测中的洞见涌现

产生科学知识是科学研究的价值所在,而科学预测作为这一过程中的关键一环,其准确性直接决定着科学知识产生的步伐。传统科学预测往往依赖科学家的理性分析、演绎推理与抽象建模。AI4S 背景下科学预测方式正经历深刻变革,通过预训练模型能够实现无需人类干预的科学预测,这一转变

不仅重构科学预测的实践路径,更为多个领域催生丰富的科学洞见。基于 AI 的科学预测,融合深度学习算法与庞大数据集的训练,结合学科领域的先验知识,构建高度专业化和智能化的预测模型。这些模型具备预测新物质、新结构或新变化的能力,能为研究者提供全新的研究视角与方向。同时,模型强大的特征分析与模式识别能力,能够高效处理大规模、高维度数据,实现自动化的数据处理与结果预测,极大提高预测效率与精度,减轻研究者的负担,提升科学研究生产力。更重要的是,这种预测方式有助于打破科研中个体的思维框架与限制,促进多学科交叉和跨学科的知识共享,激发新的研究思路与方法,助力集体智慧的涌现,推动科学发现边界的不断拓展。

一些研究案例证明了 AI 在科学预测中的巨大潜力和广阔前景。麻省理工学院的科学家利用经过约 2 500 个分子训练的 AI 模型,成功从数百万种候选化合物中预测出能够对抗“超级细菌”的新型抗生素 halicin,并通过生物实验验证了其有效性<sup>[19]</sup>。在生物分析领域,AlphaFold2 蛋白质结构解析模型能够迅速预测目标蛋白质的结构,甚至揭示自然界中尚未发现的蛋白质结构。此外,美国加州大学伯克利分校和劳伦斯国家实验室的“A-Lab”自主实验系统,展现出强大的自主发现和预测新材料的能力。当前,大模型正逐步向多模态通用性发展,研究者和开发者能够更便捷地训练自定义模型,满足不同领域科学研究的需求。加州理工学院、麻省理工学院及 Nvidia 公司,基于开源大语言模型联合构建了定理证明器,为数学领域的形式化定理证明提供了有力支持<sup>[20]</sup>。微软研究院研发的 DiG 模型框架,在处理不同类型的分子系统和描述符方面表现出色,主要用于预测分子结构平衡分布,为分子科学研究、药物设计以及材料科学等领域的研究带来新机遇<sup>[21]</sup>。IBM RXN 云服务则基于数百万个合成有机化学反应训练的 Transformer 模型,为化学反应预测和有机合成实验提供新思路。

## 二、AI 赋能高校科学研究范式创新的风险检视

数智技术无疑为高校科学研究注入了前所未有的活力,显著加速科学研究进程,拓宽科学发现视野。然而,AI 的双刃剑效应也随之显现,新技术浪潮下亟须审视 AI 引发的系列风险挑战。对高校科学研究中存在的主体工具性依赖、数据偏见导致的误导性主张、算法的复杂性与不透明性引发的解释困境,以及虚假内容生成带来的信任危机等问题,必须予以高度警觉和审视剖析。

### (一)人在“旁路”:技术依赖导致主体性缺失

AI4S 背景下,数智技术与科学研究的深度融合无疑增强了研究的客观性和严谨性。然而,这种深度融合也伴随着一种隐忧,即过度依赖 AI 技术可能导致研究者的主体性地位逐渐淡化。

一方面,研究者期待通过 AI 消除科学研究过程中难以避免的主观性、偏见与认知局限,以实现科学设计与实施的优化。然而,当这种期待转变为对 AI 的过度依赖,甚至将其视为“万能钥匙”时,人的主体性在科学研究中便逐渐丧失,从而引发一种被称为“理解幻视”的认知风险。所谓“理解幻视”,即主体错误地认为借助 AI 技术能弥补自身的认知局限,进而产生过度的自信,认为自身对研究对象的了解已经超越实际水平。这种过度依赖不仅阻碍研究者深入研究和广泛探索,更可能限制科学研究的发展空间。

另一方面,大型科学研究平台与自动化实验系统的广泛应用,确实显著提升了实验的效率,加速了科学发现的步伐。然而,在高度自动化的实验过程中,人的决策性地位逐渐削弱,不再占据科学研究的主导地位。AI 研究系统能够自动完成文献搜索、综述、研究方向生成、研究假设提出,甚至提供完整的研究设计方案,并自动化处理庞大数据集以产生研究结论。AI 在提升科学研究效率的同时,也引发科学研究本质的变化。在该模式下,AI 似乎成为科学知识创造的主体,研究者则似乎被置于科学研究的“旁路”。在此状况下,主体的批判性思维面临严峻挑战,科学创造被限于 AI 自主生成的孤岛中,导致科学创意匮乏。缺乏人类批判性思维的参与,科学研究往往难以触及核心问题,也难以

产生真正的影响力,这也是新技术背景下科学研究成果在数量上显著增长但具有颠覆性价值的研究却逐渐减少的可能原因之一<sup>[22]</sup>。诚然,数智技术在科学研究中发挥着不可或缺的作用,但它始终只能作为得力工具,人类认识与改造世界仍然离不开头脑中的“奇妙计算机”。

## (二)数据陷阱:数据偏见引发误导性主张

数据作为对事物客观描述的基石,若在收集、处理与分析的各个环节中遭遇边缘化、覆盖面不全或样本偏向性等问题,基于数据生成的结果便会陷入错误性、不完整性的境地,进而引发误导性主张危机。AI 模型训练过程的复杂性与阶段性使得数据偏见显现于多个环节。

在数据采集阶段,用于模型训练所采集的数据,其来源的局限性、种类的单一性以及标签划分的主观性,都有可能引发数据偏见。耶鲁大学的一项研究揭示了这一点,利用 AI 的皮肤病学算法在诊断不同种族人群时,准确性存在显著差异,这背后的原因,正是训练模型所采集的数据大部分来自白人,导致算法应用在其他种族时产生偏见<sup>[23]</sup>。

在模型训练阶段,训练过程的不严谨同样会埋下偏见的隐患。例如,训练数据和测试数据集划分不明确、充斥重复数据、混杂伪数据,都可能导致算法模型本身带有偏见。普林斯顿大学 Kapoor 团队指出,至少在 17 个领域和数百篇论文中存在可重复性危机,这背后正是基于伪研究数据所训练的模型,其生成的研究结果也具有误导性<sup>[24]</sup>。同时,模型训练集和测试集的混淆,也将引发拟合、功能泛化、能力不足等问题,进一步加剧科学研究的质量危机,并引发 AI 知识生成的可信度危机。

开发者的个人偏向也将对算法产生影响。由于开发者基于自身的认知偏见与情感倾向,在算法设计时可能会不自觉地强调某一特定特征而忽视其他关键变量,这种偏见性在算法的实际应用中表现为对不同用户或系统的不公平。同时,算法设计群体存在单一性特征。美国平等就业机会委员会观察到“高科技”行业中白人比例过高,这一现状将加剧算法开发中的种族歧视,使模型无形中偏向部分特定群体,并对其他群体产生歧视<sup>[25]</sup>。

## (三)算法黑箱:不透明决策导致解释困境

科学研究中的算法“黑箱”现象日益凸显,这源于技术本身的复杂性与技术开发者的算法保护策略。研究者在使用 AI 时,往往无法洞察算法生成结果背后的设计逻辑与内部机制,导致对算法结果的解释、监管与审查陷入重重困境。这种不透明性不仅影响研究结论的可靠性,也阻碍研究者对算法决策的信任。一方面,研究中算法与模型设计的科学性不足,加之算法内部复杂的数据和参数设置,使得算法的运作过程和决策逻辑难以被完全理解,研究者在使用这些模型时,可能会基于不准确的预测或误导性的结论开展研究,进而对整个人类知识库构成威胁。例如,研究人员在调整数据和参数以符合个人预期时,AI 为其提供了过度的自由度,可能导致研究结果失真。另一方面,算法的不透明性也限制了研究者难以准确追溯问题根源,不知不觉中可能陷入学术信息茧房,导致视野受限,创新受阻。此外,推荐算法机制倾向于推送与用户偏好高度一致的内容,将加剧茧房效应,限制研究者的研究视角与深度。

算法的复杂性和不透明性也将加剧科学研究内容审查的困难。一方面,由于无法直接窥视算法的内部运作,研究者往往难以对算法生成的结果进行有效评估与监管,从而难以确保研究结论的准确性和可信度。另一方面,科学研究活动中涉及数据收集、处理、分析和解释等环节,算法的不透明性会使得研究者难以确保项目的合规性和伦理性,引发数据滥用、隐私泄露等伦理问题。此外,法律法规、伦理标准以及监管机构的审查方式具有一定滞后性,无法及时应对算法带来的各种风险。以生物医学领域为例,AI 算法在图像分类中的应用虽然具有巨大潜力,但错误的分类结果可能直接关系到患者的生死。因此,确保算法的透明性和可靠性,是科学研究领域亟待解决的问题。

## (四)信任危机:大模型滋生虚假内容泛滥

AI 特别是生成式 AI 的迅猛发展,使其内容生成能力达到了前所未有的高度。然而,这种技术的

滥用却导致学术造假问题频发。尤其在高校中,学生直接利用生成式 AI 技术完成课程论文和作业,引发严重的学术诚信危机。马里兰大学的学生行为办公室在 2022—2023 学年中收到的与 AI 相关的学术诚信违规举报高达 73 起,令人震惊的数字背后,凸显高校学术诚信问题的紧迫性<sup>[26]</sup>。在科学研究领域,生成式 AI 的广泛应用与论文发表所要求的原创性声明原则之间存在明显的冲突。研究表明,当 ChatGPT 创建的摘要被提交给学术评审员时,仅有 63% 的“赝品”被发现<sup>[27]</sup>。这一发现揭示了学术造假监管工作的艰巨性,也让我们意识到,在享受技术带来便利的同时,必须警惕其潜在的风险。确保科学研究的学术公正性,对于维护科学研发生态的稳定运行至关重要。高校作为学术培养和知识生产的主要场所,理应承担起监管和治理的重任。

AI 大语言模型在科学研究中的广泛应用确实带来诸多便利,但频繁生成的虚假信息以及缺乏有效监管机制的现实,已经引发研究者对 AI 技术的信任危机。一方面,模型的训练依赖庞大的训练集,然而训练集来源的多样性、复杂性与冗余性,使得数据的真伪性与准确性难以保证。此外,人为的恶意训练与对抗攻击,也可能导致模型生成虚假、错误或误导性内容。另一方面,相较于传统实验研究,利用 AI 进行实验与研究往往缺少必要的申请、批准环节,以及缺乏如 IRB(机构审查委员会)机构的督查。科学研究中 AI 使用的规范与伦理尚未得到系统、标准的管理,其使用过程缺乏透明度和可控性,进一步加剧研究者对人工智能的信任危机。*Nature* 对 1 600 名研究人员的调查显示,超过半数的研究者担心 AI 生成的错误内容激增,认为抄袭变得更容易且难以检测,担忧其可能导致错误或不准确的研究结果,更有近六成的研究者认为这些工具可能引发欺诈行为<sup>[28]</sup>。这一担忧不容忽视。

### 三、AI 赋能高校科学研究范式创新的实践进路

科学研究范式的转型并非意味着对过往范式的彻底背离,而是要在已有范式的基础上持续创变。高校作为实现高水平科技自立自强的关键力量,应积极融入数智技术浪潮,利用技术变革为科学研究创新注入新动力。同时,需保持敏锐的洞察力,主动应对技术衍生的风险与挑战,在实践中探寻一条深植于中国学术土壤且充分展现中国独特话语体系的高校科学研究发展道路,以推动科学研究的持续繁荣与进步。

#### (一)人在回路:以跨界协同提升 AI 科研素养

高校在开展科学研究的过程中,应始终坚守人的主体地位,避免陷入“理解幻视”困境。所谓“人在回路”,指机器决策中强调人的决策参与,构建融合人类智慧的算法,通过跨界融合提升研究者的 AI 科研素养,并设计相应的培训方案,确保“科学研究—人才培养”良性循环和可持续发展。

其一,开发以人为本的智能科研系统,构建人在回路的 AI 算法。这一系统强调人机协同,其中人的作用与地位至关重要,AI 作为辅助者而非替代者。系统需要具备高度的可干预性和适应性,使研究者能随时干预和调整系统运行,确保科学研究方向的正确性。同时,算法的设计应具备透明性、可解释性,确保研究者能够清晰理解算法的运行过程和结果,从而进行准确的评估。以华盛顿大学研究团队为例,他们在人机协同科学研究过程中,充分发挥人的主导作用,结合 AI 技术提议,采用扩展 RFDiffusion 与序列设计工具 ProteinMPNN 结合的新方法,充分发挥了实验设计中人类思维的创新性<sup>[29]</sup>。

其二,倡导各界紧密合作,跨界融合培育 AI 科研素养。通过建立“高校—政府—企业”三位一体的跨界协作平台,构建“AI+X”跨专业融合课程体系,推动 AI 技术与其他学科领域的深度融合,全面加强研究者 AI 科研素养的培养<sup>[30]</sup>。譬如,麻省理工学院联合 8 个不同院系,依托“计算机科学与 AI 实验室”,为不同需求层次的学生提供多样化的学位项目(包括计算与认知工程硕士、计算机科学博士、计算机科学与工程博士等),并积极与政府、学术研究机构等合作,为学生提供丰富的实习机会<sup>[31]</sup>。帝国理工学院联合康奈尔大学以及政府、行业,建立了跨大西洋 AI 科学网络,发布 I-X 创新

计划,将关键学科(如材料、物理学、生物学或可持续科学领域)的 AI 科学家聚集起来,共同探索该技术在科学工程研究中的应用,以加速科学发现<sup>[32]</sup>。

其三,依托跨界共同体,构建学分认定与资金保障机制,为跨学科融合培育提供支持。在学分互认方面,可以为科学研究后备军提供灵活多样的学习路径和成果认证方式。如我国华东地区 6 所高校联合华为、百度、商汤科技等企业,采用共建共选的方式,以小规模限制性在线课程(SPOC)的形式,开设了“AI+X 微专业”,学生在 1-2 年内获得不少于 12 个学分,就可被授予由 6 所高校共同签章的微专业证书,学分可以申请转换为校内公共选修课学分<sup>[33]</sup>。在资金支持与激励方面,Schmidt Futures 项目面向高校设立了高达 1.48 亿美元的 AI 科学奖学金,培养未来的 AI 领导者,目前已在促进数学发现、监测作物干旱、改进太阳能材料等多个领域取得显著成效<sup>[32]</sup>。

## (二)数据规范:以数据标准确保合规使用

AI 模型的训练和学习过程离不开数据的反复训练与测试,数据的客观性、中立性、全面性,数据库的标准化程度,都将影响 AI 算法决策的结果。若数据集在划分过程中缺乏严谨性,还可能导致模型偏见问题。因此,高校科学研究应致力于数据规范化,构建清晰的数据清单和统一的数据标准框架。

首先,明确数据标准,建立详尽数据清单。通过系统化数据存储与精细化管理,确保数据的准确性、完整性和一致性。同时,遵循统一的标准进行数据的标准化操作,明确数据的来源、种类和标签的划分。目前,FAIR 标准(即可查找、可访问、可互操作和可重复使用)已在多个领域得到广泛认可,如跨学科地球数据联盟 EarthChem 图书馆,围绕业界标准整合数据集格式,与不同学科领域保持高度一致,并能在互联网上便捷地获取<sup>[34]28</sup>。

其次,基于标准化数据,构建高质量数据集和数据库。建议从国家层面推动跨行业、高校、企业以及科学研究机构的合作,共同构建超大型数据集和数据库,为 AI 技术的发展提供坚实的数据支撑。在国际方面,谷歌联合哈佛大学、麻省理工学院等高校,依据统一数据标准整合处理了包含 25 000 TB 小鼠大脑特定区域图谱数据集,为相关研究提供宝贵数据源<sup>[35]</sup>。在国内,腾讯优图实验室联合多个部门,基于统一标准建立了全球最大的甲骨文单字数据库,为后续相关研究工具的开发提供支持<sup>[36]</sup>。

再次,明晰数据使用标准,规范模型训练过程。模型训练过程中操作不当,将导致数据偏见等问题,进而影响 AI 模型的准确性。因此,需要建立相对统一的标准,包括数据使用、模型训练策略选择、评估与反馈指标等。国内外众多机构已发布数据使用标准和指南,如 FAIR 原则中对数据在代码中的存放、研究论文的引用进行了规范<sup>[34]26</sup>;欧盟委员会 FAIR 数据专家组的报告对数据的引用所需步骤进行了罗列与说明<sup>[34]26</sup>;澳大利亚研究数据共享组织(ARDC)制作了在线培训指南,引导研究人员正确引用数据、样本和软件资源<sup>[37]</sup>;美国研究机构发布了如何增加科学数据访问权限的指南,以规范数据的使用<sup>[38]</sup>;斯坦福大学、圣地亚哥州立大学等高校发布相关生成式 AI 教学指南与使用指南,分别对教师、学生等不同群体提供使用手册,并围绕使用的过程步骤进行指导与规范<sup>[39]</sup>。

## (三)可解释 AI:以透明算法研发解密“黑箱”

算法模型在数据处理中直接输出结果,其内部运行机制复杂,即使对算法工程师或数据科学家而言也是一片迷雾,这种不透明性极大地削弱了决策者对 AI 技术的信任。为应对“黑箱”挑战,可解释 AI(XAI)应运而生。XAI 系统致力于揭示 AI 模型的内部逻辑,详细解释 AI 决策和预测的依据与过程,揭示潜在的影响和偏差,并展现模型的准确性、公平性、透明度和结果。其不仅为用户提供理解和信任算法结果的途径,也支持技术人员以负责任的态度进行开发,同时赋予决策者质疑 AI 结果并据此作出明智决策的能力。在推动 AI 变革科学研究范式的过程中,高校科学研究应积极采用 XAI 工具,以提高算法模型的透明度,帮助研究者和 AI 开发者更好地理解 and 信任新一代的 AI 科研伙伴。

在高校科学研究领域,XAI 的价值已得到广泛认可。加州理工学院在虚拟研讨会上,明确将开发

XAI 列为重要议题,强调 AI 模型应具备精确的开发校准、可验证的单元测试以及用户友好的界面,确保实际应用中的高效性和可靠性<sup>[40]</sup>;UCLA 的研究团队利用 XAI 技术对原本用于山体滑坡预测的深度神经网络(DNN)进行优化,成功开发出 SNN 模型<sup>[41]</sup>,不仅改善了 DNN 在解释性方面的不足,还保持了高准确性、高泛化能力和低模型复杂度,为山体滑坡的精准预测提供有力支持。在国内,山东大学开发的 RetroExplainer 算法在有机物的逆合成路线识别中取得显著进展,不仅提高了识别效率,还为用户提供了对算法决策过程的深入理解<sup>[42]</sup>;清华大学的研究团队利用可解释的机器学习技术对 BiVO<sub>4</sub> 光阳极系统进行优化,利用 AI 模型分析过去的实验数据,揭示了系统内部的复杂关系,为进一步优化提供重要指导<sup>[43]</sup>。目前,在科学研究中用于缓解算法“黑箱”问题的 XAI 主要有两类实现路径。一类是设计本质上可解释的预测模型,这类模型通过采用基于规则、决策树和线性模型等特征权重、内部路径和规则可见的组件,确保决策过程的可追溯性和透明度<sup>[44]</sup>。然而这种方法也面临功能和预测性上的限制,因为基于规则的模型组件往往较为固定,用户只能进行微调。另一类是事后可解释性技术,主要解决深度学习模型在追求预测准确性时牺牲透明度和可解释性的问题<sup>[45]</sup>,包括全局模型和局部模型两个维度,前者揭示黑盒模型的平均行为,后者则针对单个预测进行解释。

#### (四) 伦理审查:以科研治理促进可持续发展

AI 赋能科学研究的同时,引发了虚假信息生成和学术造假问题,这导致科学研究领域对 AI 的信任度下降。因此,对 AI 进行伦理审查与治理已成为当务之急。

首先,积极开发用于识别造假的 AI 工具,利用 AI 技术对造假问题及虚假信息进行规范与治理,并不断提升治理效率和准确性。政府应给予技术创新研究资金支持及资源激励,确保技术的可持续发展。例如,OpenAI 的 Deepfake 探测器能有效检测 DALL-E 等图像生成器创建的虚假内容<sup>[46]</sup>;百度 AI 开放平台推出了图片造假审查工具,能够精准识别图片是否经过篡改;哈佛大学在医疗 AI 领域投入巨资,集结顶尖科学研究人才,实现了技术创新突破<sup>[47]</sup>。

其次,设立大模型伦理委员会,统筹各部门资源共同应对 AI 伦理风险。应汇聚各行业领域的共治力量,形成治理共同体,并重视社会中每个成员的力量,使其具有对技术限制的发言权。正如此勒陀利亚大学数据科学主席 Marivate 所言,AI 治理是一项团队活动,道德决策和责任应由全社会共同承担<sup>[48]</sup>。

再次,尝试构建全面 AI 科学研究治理框架,阐明治理体系、机制和条例,鼓励公众参与治理,提供具体参与路径,通过培训研讨、课程开发等途径,推广 AI 研究工具,普及应用伦理规范,制定 AI 研究工具的使用原则、规定与指南,从而推动治理流程的高效运转<sup>[49]</sup>。在框架明晰的基础上,应进一步细化实施策略,例如设立专门的治理机构,严格监测大模型输出,确保输出内容的真实性与准确性;加强算法的备案和变更管理,对算法的合法性、合规性和安全性进行全面审查与评估。此外,政府、科研机构、企业等各方应密切合作,共同明确治理流程、框架及实施路径。英国政府在数据伦理方面树立了典范,不仅制定了详细的数据伦理框架(涵盖内部监测与治理机制、数据处理责任、数据处理活动记录等方面),还明确了数据处理中的责任,提出公众问责制,确保公众或其代表能够有效监督政府决策与行动<sup>[50]</sup>。

## 四、结 语

从寻找数据模式的机器学习模型,到从海量文本代码中生成内容的最新通用算法,人工智能不仅加速了知识探索的步伐,更为科学探索开辟了新航道。在人工智能赋能的新科学研究工作模式下,科学研究工作者的想象力不断被激发,能够洞察潜在的数据模式与异常现象,跨越自身的认知边界,对已有研究作出新解释,并发现更具价值的新问题。人工智能赋能科学研究改变了工作的流程,引领科学研究范式的创变。新科学研究范式中人工智能引领知识发现,成为人类创造力的强大催化剂,依托数据驱动

的预测模型,融合模拟技术与可扩展计算的力量,激发了研究者的创新思维,形成创新性的解决方案。

新局面意味着新风险,高校作为科学研究的主阵地,必须以高度的责任感与前瞻性的战略眼光进行全局性部署,确保人工智能在科学研究中可靠运行,同时采取周密的伦理审查措施。如此,才能实现人工智能益处最大化、潜在风险最小化,解锁以前无法挖掘的科学奥秘,推动人类社会迈向更加辉煌的未来。展望未来,人工智能与科学研究的深度融合将引领一系列前沿方向,涵盖从方法论的革新到自主科学研究系统的构建,再到知识提取、复杂现象建模、假设生成验证及新学科建设与跨学科合作的全方位推进,在提升数据质量、优化实验设计、挖掘潜在模式等方面发挥关键作用,在新兴科学领域人工智能更进一步展现巨大潜力,为科学研究开辟新道路。

### 参考文献:

- [1] 李国杰. 智能化科研(AI4R):第五科研范式[J]. 中国科学院院刊,2024,39(1):49.
- [2] 张海生. 人工智能赋能学科建设:解释模型与逻辑解构[J]. 高校教育管理,2023,17(3):4250,75.
- [3] 库恩. 科学革命的结构[M]. 金吾伦,胡新和,译. 北京:北京大学出版社,2003:9.
- [4] 张海生. 人工智能赋能大学治理:多重效应与治理效能转化[J]. 重庆高教研究,2024,12(2):2536.
- [5] 亨普尔. 自然科学的哲学[M]. 张华夏,余谋昌,鲁旭东,译. 北京:中国人民大学出版社,1986:12.
- [6] 周林东. 科学哲学[M]. 上海:复旦大学出版社,2004:128.
- [7] ZHOU C, PAFFENROTH R C. Anomaly detection with robust deep autoencoders[C]//Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining. New York: Association for Computing Machinery,2017:665674.
- [8] WANG H C, FU T F, DU Y Q, et al. Scientific discovery in the age of artificial intelligence[J]. Nature,2023,620(7972):4760.
- [9] How AI is transforming scientific research, with Rebecca Williett(EP. 117)[EB/OL]. (20230810)[20240618]. <https://news.uchicago.edu/how-ai-transforming-scientific-research>.
- [10] GOVORKOVA E, PULJAK E, AARRESTAD T, et al. Autoencoders on field-programmable gate arrays for real-time, unsupervised new physics detection at 40 MHz at the Large Hadron Collider[J]. Nature machine intelligence,2022,4(2):154161.
- [11] ZHOU Y K, CHIA M A, WAGNER S K, et al. A foundation model for generalizable disease detection from retinal images[J]. Nature,2023,622(7981):156163.
- [12] TSHITOYAN V, DAGDELEN J, WESTON L, et al. Unsupervised word embeddings capture latent knowledge from materials science literature[J]. Nature,2019,571(7763):9598.
- [13] SOURATI J, EVANS J A. Accelerating science with human-aware artificial intelligence[J]. Nature human behaviour, 2023,7(10):16821696.
- [14] REMI L, ALVARO S G, MATTHEW W, et al. Learning skillful medium-range global weather forecasting[J]. Science, 2023,382(6677):14161421.
- [15] DUAN C R, DU Y Q, JIA H J, et al. Accurate transition state generation with an object-aware equivariant elementary reaction diffusion model[J]. Nature computational science,2023,3(12):10451055.
- [16] COLE A. Four Ways AI Has Already Changed Science[EB/OL]. (20230718)[20240618]. <https://www.techopedia.com/four-ways-ai-has-already-changed-science>.
- [17] 王敏.“机器化学家”是怎样炼成的[N]. 中国科学报,20230516(01).
- [18] The Economist. Could AI transform science itself?[EB/OL]. (20230913)[20240618]. <https://www.economist.com/science-and-technology/2023/09/13/could-ai-transform-science-itself>.
- [19] STOKES J M, YANG K, SWANSON K, et al. A deep learning approach to antibiotic discovery[J]. Cell, 2020,181(2):475483.
- [20] YANG K, SWOPE A M, GU A, et al. Leandjo:theorem proving with retrieval-augmented language models[EB/OL]. (20231016)[20240618]. <https://arxiv.org/abs/2306.15626>.
- [21] ZHENG S X, HE J Y, LIU C, et al. Predicting equilibrium distributions for molecular systems with deep learning[J].

- Nature machine intelligence, 2024,6(5):558567.
- [22] PARK M, LEAHEY E, FUNK R J. Papers and patents are becoming less disruptive over time[J]. Nature,2023,613(7942):138144.
- [23] ZACK T, LEHMEN E, SUZGUN M, et al. Assessing the potential of GPT-4 to perpetuate racial and gender biases in health care: a model evaluation study[J]. The Lancet digital health,2024,6(1):e12e22.
- [24] Nature. AI will transform science—now researchers must tame it [EB/OL]. (2023-09-27) [2024-06-18]. <https://www.nature.com/articles/d41586023029886>.
- [25] SOLOMON D, MAXWELL C, CASTRO A. Systematic Inequality and Economic Opportunity[EB/OL]. (2019-08-07) [2024-06-18]. <https://www.americanprogress.org/article/systematic-inequality-economic-opportunity/>.
- [26] GAUR A. UMD records at least 50 AI-related academic integrity cases in 2022-23[EB/OL]. (2023-11-27) [2024-06-18]. <https://dbknews.com/2023/11/27/umd-chatgpt-academic-integrity-cases/>.
- [27] HOLDEN T H. ChatGPT is fun, but not an author[J]. Science,2023,379(6630):313.
- [28] VAN N R, PERKEL J M. AI and science: what 1,600 researchers think[J]. Nature,2023,621(7980):672675.
- [29] RUMFIEL G. New proteins, better batteries: Scientists are using AI to speed up discoveries[EB/OL]. (2023-10-12) [2024-06-18]. <https://www.npr.org/sections/health-shots/2023/10/12/1205201928/artificial-intelligence-ai-scientific-discoveries-proteins-drugs-solar>.
- [30] 祝智庭,戴岭,赵晓伟,等. 新质人才培养:数智时代教育的新使命[J]. 电化教育研究,2024,45(1):5260.
- [31] 李锋亮,庞雅然. 世界一流大学如何建设人工智能学科[N]. 光明日报,2022-02-10(14).
- [32] JOHNS S. Imperial and Cornell University to cooperate on AI in scientific discovery[EB/OL]. (2023-04-18) [2024-06-18]. <https://www.imperial.ac.uk/news/244392/imperial-cornell-university-cooperate-ai-scientific/>.
- [33] 吴飞,陈为,孙凌云,等. 以知识点为中心建设 AI+X 微专业[J]. 科教发展研究,2023,3(1):96116.
- [34] European Commission, Directorate-General for Research and Innovation. Turning FAIR into reality: final report and action plan from the European Commission expert group on FAIR data[R]. Luxembourg: Publications Office of the European Union,2018.
- [35] JANUSZEWSKI M. Google Research embarks on effort to map a mouse brain[EB/OL]. (2023-09-26) [2024-06-18]. <https://research.google/blog/google-research-embarks-on-effort-to-map-a-mouse-brain/>.
- [36] 腾讯 AI 新成果:建立全球最大的甲骨文单字数据库 [EB/OL]. (2023-07-25) [2024-06-18]. <https://cloud.tencent.com/developer/article/2305113>.
- [37] ARDC. Citation and Identifiers [EB/OL]. (2022-05-14) [2024-06-18]. <https://ardc.edu.au/resource/citation-and-identifiers/>.
- [38] HOLDREN J P. Memorandum for the Heads of Executive Departments and Agencies: Increasing Access to the Results of Federally Funded Scientific Research[EB/OL]. (2013-02-22) [2024-06-18]. <https://rosap.ntl.bts.gov/view/dot/34953>.
- [39] Artificial Intelligence Teaching Guide[EB/OL]. (2024-06-18) [2024-06-18]. <https://teachingcommons.stanford.edu/teaching-guides/artificial-intelligence-teaching-guide>.
- [40] Caltech Explainable AI Virtual Workshop [EB/OL]. (2021-09-23) [2024-06-18]. <https://sites.astro.caltech.edu/xai4s/program.html>.
- [41] UCLA. UCLA team uses artificial intelligence to predict landslides—UCLA [EB/OL]. (2024-06-18) [2024-06-29]. <https://www.chemistry.ucla.edu/news/ucla-team-uses-artificial-intelligence-to-predict-landslides/>.
- [42] WANG Y, PANG C, WANG Y Z, et al. Retrosynthesis prediction with an interpretable deep-learning framework based on molecular assembly tasks[J]. Nature communications,2023,14(1):6155.
- [43] HUANG M, WANG S, ZHU H. A comprehensive machine learning strategy for designing high-performance photoanode catalysts[J]. Journal of materials chemistry, A. materials for energy and sustainability,2023,11(40):2161921627.
- [44] RAI A. Explainable ai:from black box to glass box[J]. Journal of the academy of marketing science:official publication of the academy of marketing science,2020,48(1):137141.
- [45] MARGOT V. A Brief Overview of Methods to Explain AI (XAI) [EB/OL]. (2024-11-27) [2024-06-18]. <https://towardsdatascience.com/a-brief-overview-of-methods-to-explain-ai-xai-fe0d2a7b05d6>.
- [46] METZ C, HSU T. OpenAI Releases “Deepfake” Detector to Disinformation Researchers [EB/OL]. (2024-05-07)

- [20240618]. <https://www.nytimes.com/2024/05/07/technology/openai-deepfake-detector.html>.
- [47] Artificial Intelligence in Medicine Program[EB/OL]. [20240618]. <https://aim.hms.harvard.edu/>.
- [48] FRUEH S. How AI Is Shaping Scientific Discovery[EB/OL]. (20231106) [20240618]. <https://www.nationalacademies.org/news/2023/11/how-ai-is-shaping-scientific-discovery>.
- [49] 常桐善,赵蕾. 美国高校应对和使用人工智能工具的策略与原则[J]. 重庆高教研究,2024,12(4):6879.
- [50] Central Digital and Data Office. Data Ethics Framework[EB/OL]. (20200916) [20240618]. <https://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework-2020>.

(责任编辑:杨慷慨 校对:吴朝平)

## Empowering Innovation in Scientific Research Paradigms at Higher Education Institutions with Artificial Intelligence: Value, Risks and Approaches

ZHAO Xiaowei, WANG Xiaoyu, WANG Yirong, SHEN Shusheng  
(College of Educational Sciences, Nanjing Normal University, Nanjing 210097, China)

**Abstract:** Colleges and universities are the main battlefield of scientific research. As the main force of basic research and major scientific and technological breakthroughs, university researchers shoulder the important mission of accelerating the formation of new quality productivity. Under the wave of the new round of scientific and technological revolution, artificial intelligence technology is widely applied and deeply integrated in various aspects of scientific research, not only reshaping the basic process of scientific research, but also leading the transformation of scientific research paradigm to the intelligent science research paradigm. From the perspective of logical positivism, exploring the multiple values of AI empowering innovation in scientific research paradigms in universities, including enhancing data collection in scientific research, accelerating the generation and verification of scientific hypotheses, achieving experimental simulation and automation, and stimulating scientific insights and creativity, can help understand the value and profound impact of AI empowering scientific research paradigms. Based on this, the challenges and risks encountered in the current transformation of scientific research paradigms were examined, including the risk of human marginalization (the phenomenon of “human-on-the-side”), the dual dilemma of data bias and misleading claims (the “data trap” issue), the lack of algorithm transparency (the “algorithm black box” issue), and the resulting crisis of trust. Universities should closely rely on the new round of technological changes in the transformation of scientific research paradigms, root themselves in the soil of Chinese scientific research practice, seek innovation based on existing paradigms, and explore a scientific research path that is in line with China’s national conditions and has distinct Chinese characteristics. Specifically, it is necessary to strengthen interdisciplinary cooperation and collaborative innovation, standardize data usage standards, promote the development of interpretable AI, and establish and improve research ethics review mechanisms, to provide solid theoretical support and action guidelines for innovative research paradigms.

**Key words:** scientific research paradigm; intelligent scientific research paradigm; artificial intelligence; data trap; crisis of confidence