

大部分孩子都有长大后成为科学家的梦。那么，人工智能科学家是如何成长的？他们在做哪些有趣的科学研究？如何成长为像他们一样的科学家？为了让更多的读者了解人工智能科学家，来自温州科技高中的特级教师谢作如特邀上海人工智能实验室青年科学家白磊博士进行对话。白磊博士从事人工智能与地球科学领域的交叉研究工作，负责的“风乌”模型实现了10公里分辨率的全球气象预报，是目前分辨率最高的人工智能气象大模型。

# AI科学家眼里的青少年AI教育



» 谢作如

温州科技高中人工智能科创中心负责人，浙江省特级教师，中国电子学会现代教育技术分会副主任委员，上海人工智能实验室XEdu项目负责人，OpenHydra项目发起人。



» 白磊

上海人工智能实验室青年科学家，新南威尔士大学博士，悉尼大学博士后研究员，曾获2022年世界人工智能大会云帆奖、2020年新南威尔士大学工程研究卓越奖、2019年谷歌博士奖学金等。

## 科研与成长经历

**谢作如：**随着生成式人工智能的发展，中小学AI教育再次引发热议。感谢白磊博士在百忙之中接受邀请，聊一聊“青少年AI教育”这一热点话题。首先，请简要介绍一下您在上海人工智能实验室的工作以及研究成果。

**白磊：**感谢谢老师邀请。我在实验室主要负责人工智能与地球科学领域的交叉研究工作，开发人工智能算法来对围绕咱们地球的大气、海洋、地质等圈层建模，

从而实现对地球系统的实时高分辨率监测和面向不同时空尺度的高精度高效率预报，并进一步服务气象预报、海洋预报、能源预测与平衡等国家和社会需求。我最近的研究成果是“风乌”气象大模型系列，2023年4月我们联合中科大、上交大、南信大、上海中心气象台等高校和机构发布了“风乌”气象大模型，首次将全球中期气象预报的可用时效提高到10.75天，模型实验部署于国家气象中心、上海市气象局、香港天文台等机构，助力我国

做科研工作需要对这个世界或者所做的事情时刻保有好奇心和求知心，希望把一件事弄明白或者做好，然后才能全身心地投入精力去做。

的气象预报。今年3月，我们又发布了“风乌”GHR版本，首次利用人工智能算法实现了10公里分辨率的全球气象预报，是目前分辨率最高的人工智能气象大模型。

**谢作如：**能否介绍一下您的求学经历，您的成长经历与现在从事的科研工作有什么联系？

**白磊：**我博士就读于新南威尔士大学，当时的主要研究方向是时空预测学习，通过开发人工智能算法实现对不同地理位置上持续收集的多维时间序列流进行预测。博士毕业后，我在悉尼大学进行博士后研究，主要探索大规模通用预训练技术。这些工作为我探索人工智能驱动的地球科学提供了很好的支撑，因为地球科学领域的大部分数据类型都是多源的时空数据，而且数据规模很大，我之前的研究经历刚好可以满足这方面

的需求。

**谢作如：**您在中小学是否接触过科研工作？或者说，在中小学阶段有哪些经历对您现在从事的工作产生影响？

**白磊：**我在中小学阶段没有接触过科研工作。至于说有联系的经历，我想首先是整个中小学的学习生活培养了我求知和比较乐观的精神，这是从事科研工作必需具备的。做科研工作需要对这个世界或者所做的事情时刻保有好奇心和求知心，希望把一件事弄明白或者做好，然后才能全身心地投入精力去做。同时，科研工作的日常又往往是充满了失败的，绝大多数的科研成果应该都是在无数的失败尝试后取得的，因此需要科研工作者自身有乐观的精神，不断失败，不断尝试。

## AI4S：人工智能如何赋能科学研究

**谢作如：**我是在2023年通过新闻才了解到实验室有个“风乌”大模型团队，并在预报方面实现了重大突破。“风乌”取名自秦汉时期的“相风铜乌”，这是世界上最早的测风设备，仅从名字看就很有文化的传承感。您负责的“风乌”大模型是如何实现准确预报天气的？在哪些方面进行了突破？

**白磊：**“风乌”基于人工智能方法开发，借助神经网络在建模非线性关系方面的优势和大模型海量参数带来的强大拟合能力，可以更好地从历史大气数据中总结出大气系统的底层规律。从人工智能具体方法来说，“风乌”具有三个方面的创新：一是提出基于多模态大模型思想的Transformer网络设计来表征多种多样的大气变量，实现对全球高分辨率高维大气数据的高效

建模；二是提出新的优化目标，针对不同区域不同大气变量高度耦合但分布差异大的问题，提出不确定性损失函数，通过学习自动调整不同位置不同变量的优化权重，提高网络优化效果；三是针对长期预测这一难题专门设计方法，针对全球中期预报的长时序生成问题，提出“缓存回放”策略，使模型有意识地处理误差累积问题，在有限硬件的基础上间接实现了对长期预测误差的优化，显著降低了模型的长期预测误差。实验结果表明，“风乌”的10天预报误差比此前DeepMind发布的GraphCast模型降低10.87%，在“风乌”报告的880个预测指标中，有80%的准确性高于GraphCast。

**谢作如：**近几年，AI for Science（科学智能，也称AI4S）成为备受关注的科研领域。我在很多新闻中看到

这样一句话——新一代人工智能技术推动科研范式变革。什么是“科研范式”？它又是如何因为人工智能而发生变革？请您结合典型的案例介绍一下。

**白磊：**关于“科研范式”可能不同的研究者有不同的理解，在我看来，科研范式指的是大部分科研工作者采用的研究方法、研究路径。我个人认为科研范式的核心是由观测数据和建模方法两部分组成，在以前可获得观测数据非常有限的情况下，主流的建模方法需要基于定性推论或者数理模拟的方法进行，在准确性或者

效率上往往不尽如人意，而且由于需要具备非常强的专家知识，往往发展缓慢。基于人工智能的科研范式，建立在过去科学发展所积累的海量观测数据的基础上。由于数据的日益丰富和人工智能方法的相对成熟，科研工作者们有机会利用人工智能自动地对科学数据进行处理、建模，从而发现规律进行预测。基于人工智能的科研范式已经在蛋白质折叠（如AlphaFold）和气象预报（如风乌）等非常核心的科研任务上取得了成功。

## 人工智能对中小学教育的影响

**谢作如：**我记得数学家莱布尼茨曾说过，让一些杰出的人才像奴隶般地把时间浪费在计算上是不值得的。因此，才有一批科学家前赴后继，最终发明了计算机。在我上学的时候，数学教材中还有“珠算”“对数表”之类的学习内容。随着计算机的普及，这些内容慢慢被取消了。从您的角度看，随着人工智能的发展，中小学的哪些学习内容可能会随之发生变化呢？

**白磊：**这个问题很好。前两天我刚好看到美国有大学教授在呼吁“人工智能时代需要加强数学的教育”，并得到了Sam Altman、Elon Musk的支持，这说明世界上已经有许多人意识到随着人工智能的发展，我们现有的教育体系也会有相应的变化。第一，人工智能的概念和计算机科学的基础会进入到中小学的教育中，让同学们可以更早了解计算机和人工智能技术的工作原理，并且能够做一些简单的编程实践使用人工智能工具。第二，数学理论依旧非常重要，并且为了适应计算机和人工智能技术进入中小学教育，可能需要有一些针对性的调整。第三，培养学生文学性、艺术性、创造性的能力会更加重要。随着人工智能的发展，随着一些重复性的工作被AI所替代，对人类艺术性和创造性的要求会更高。

**谢作如：**面对人工智能的快速发展，有些人很焦虑，认为要早点学习不能输在起跑线上，也有人表示完全躺平，认为人工智能不过是“纸老虎”，只要学好数理

化，打好基础就行。对此您是怎么看的？

**白磊：**我想这种态度的差别对于任何新生的技术或者事物都是一样的。一方面，对于新的技术，我们永远都应该保持一种好奇心和开放性的态度，多去了解它、认识它、学习它，并使用它好的一面。另一方面，无论技术如何变化，扎实的数理基础、独立的思考能力、出色的批判精神、优秀的创造思维都应该是一个人从小学开始培养并在后续的学习中持续提高的，它们是我们的基础能力，也是我们快速学习与接纳新技术的基础。

**谢作如：**由于生成式人工智能的快速发展，有人认为中小學生只要会用人工智能（大模型）即可，会用引导词即可，甚至有人提出最重要的学科是语文，会准确描述即可。对于这些观点，您是如何评价的？

**白磊：**在前面的回答中我已经强调了文化和艺术在未来中小学教育中的重要性。此外，我想再补充两个角度来说明人工智能时代其他学科的重要性。第一个角度，当前的生成式人工智能或者说大模型还没有达到所谓的通用人工智能的程度。尽管它已经在对话、文献处理、编程、视频创作等非常多的任务上展现出了卓越的性能和提高潜力，但是它还不能完美地完成这些任务。在面对一些极端样本的时候，人工智能经常会犯错或者性能往往表现不佳。因此，我们还需要不断地开发更强大的人工智能模型、更强大的计算设备来进一步推动

无论技术如何变化,扎实的数理基础、独立的思考能力、出色的批判精神、优秀的创造思维都应该是一个人从小学开始培养并在后续的学习中持续提高的。

人工智能技术的发展,这需要我们的教育系统从小学阶段就开始培养学生的数理基础、创造性思维等。第二个角度,现实世界中的许多挑战往往需要构建复杂的大规模系统来解决,人工智能技术目前还只是系统中的一环,无法形成完成自主运行的体系,如如何自主地设计

一个航天飞机、如何构建一个完全自主运行的能源系统、如何模拟地球系统的运行进而预测天气和地震等。这些系统是我们社会日常运行的基础,也需要从小开始培养各种不同学科的高端人才来完成。

### 中小学人工智能教育和科创活动

**谢作如:**有媒体称2023年是人工智能平民化元年,因为在这一年,人工智能技术的应用越来越普及,不再是只有专业人士才能触及的技术。但在我们教育团队看来,仅仅应用人工智能显然不是在培养未来的人工智能人才。为此,我们设计了相关工具,以期降低青少年学习人工智能的门槛。借助这些工具,只要打开浏览器,用少量的代码就能训练机器学习和深度学习模型,也能用收集数据→选择SOTA模型→训练模型的方式解决一些简单的真实问题。但也有人认为,中小学没有数学的基础,即使训练了模型也不理解底层原理,并没有什么价值。您觉得中小学生学习有没有必要经历这种智能“从无到有”的过程?有没有价值?

**白磊:**我想人工智能或者说深度学习不仅仅是一种技术,也是一种科学思维和解决问题的方式。在这种体系下面,我们尝试开发新的算法来自动地从数据中识别规律,制订出解决问题的方案,并通过科学可量化的评估方式来不断改进我们的算法。虽然中小学生学习还没有扎实的数学基础来完全理解机器学习和深度学习模型的底层原理,但通过实践操作,他们可以初步理解人工智能的基本概念、科学的思维方式和用人工智能

解决问题的方式,如如何收集关于某个问题的数据、如何设计合理的评估方式来定量地评估模型的性能等。这些概念和思维方式对于做任何事情都是非常有用的。

**谢作如:**我曾经想根据AI4S的案例,设计一些面向中小学的案例,但最终因为这些案例都需要一定的学科背景知识,一直找不到好的方向。如果要为中小学生学习推荐或者设计一个人工智能的活动,您会选择怎样的主题?

**白磊:**如果单纯地是为了让中小学生学习更好地了解人工智能,培养他们对人工智能的兴趣,我想会是关于艺术和创造性方面的活动,如利用现在的大语言模型或者多模态模型进行诗歌创作和剧本写作、利用多模态生成大模型进行微电影的创作。这种活动本身非常有趣,需要的人工智能基础和学科背景知识也比较少,具有非常好的普适性,可以让更多的同学参与进来,了解人工智能技术现有的能力以及在这些任务上可能的不足。如果是希望培养同学们的人工智能技术或者是培养他们利用人工智能解决更加复杂任务的能力,那么我可能会选择把这个活动与一个具体的课程相结合(如化学或者地理),这样可以更好地根据这个课程



的实际进度和需要来设计人工智能活动的方案。

**谢作如：**我一直有个困惑，同样是“算法”一词，计算机科学中的算法比赛和人工智能中的算法比赛考核的内容是不一样的，如信息学竞赛（ACM）和人工智能算法挑战赛（Kaggle），需要的能力和提供的代码几乎是不一样的，那二者有什么区别？

**白磊：**二者的区别就类似于人工智能两种研究范式之间的区别。2019年辛顿在演讲中梳理了两种人工智能研究范式，一种是基于逻辑启发范式，即符号主义，另一种是基于生物学启发范式，即联结主义。前者认为智能是人为设计出来的，而后者认为智能因学习而来。

**谢作如：**感谢白磊博士，我们很期待像您一样的人工智能科学家能关注中小学教育，设计一些AI4S的案例。最后，请您用一段话来鼓励一下正在就读中小学的同学们吧。

**白磊：**亲爱的弟弟妹妹们，非常羡慕你们处在这样一个美好的、充满无限可能的年纪，每一天都在学习新知识，探索这个世界。虽然科幻电影在说人工智能可能会掌控世界，但是希望你们不要害怕挑战，不要畏惧困难，保持好奇，努力学习和尝试，为你们自己和我们国家的未来打下坚实的基础。同时，我更希望你们可以经常锻炼身体，经常玩耍，因为乐观的心态和强健的体魄是一切挑战的基础。



#### 对话印象——

我和白磊博士平时都在上海徐汇国际传媒港的L1大楼工作，我在12楼而白磊博士在10楼。第一次了解到白磊博士是在2022年的暑假，他做了一个关于人工智能和地球科学方向的讲座。而我近几年的研究重点就是寻找人工智能结合科创的项目，让中小学生学习AI4S领域，有空就会去听上海人工智能实验室科学家的讲座。当时，白博士的讲座给了我很大的启发。

《中国信息技术教育》杂志邀请我做几期关于人工智能教育的对话，希望能邀请人工智能领域的专家，我第一时间就想到了他。这次“对话”我们是通过共享文档软件来展开的，我提问他回答，交流非常顺畅，完稿后我们约在五楼聊天。

白博士很年轻，也很健谈，知识渊博且见解独到。关于计算机科学中的算法和人工智能的算法区别的回答解开了我的疑惑。例如，现有论文中对计算思维的界定，基本上还局限在前者，看起来需要更新观念并加入后者的内容，才能适用于当前发展现状，即智能计算时代。

人工智能的人才培养离不开基础教育，期待有更多的人工智能科学家和白博士一样，为中小学人工智能教育提供更多的智慧和力量。e