



机器人“小来”在做实验。

中国科大供图

在人们的传统印象里,在实验室搞研究的都是穿白大褂的科研人员。但在中国科学技术大学(以下简称中国科大)机器人化学家实验室里,却是另一番景象:这里没有一个人,只有动作灵活的机器人在操作台之间来回穿梭,伸出机械手臂配制试剂。它不仅能够做实验,还会自主设计实验步骤和方案。

这款全球首个集阅读文献、设计实验、自主优化等功能于一体,覆盖化学品开发全流程的机器人化学家平台,被科研人员形象地称为“机器人化学家”。

实验室负责人、中国科大化学物理系教授江俊介绍:“从数百万种材料的可能

组合中找到最优解,科研人员也许一生都做不完;有了机器人化学家,可能只需要一两周时间。”

机器人化学家融合了大数据、人工智能和自动化的强大优势,为化学这一传统学科提供了新的研究范式。

那么,这款如此聪慧的机器人化学家是如何诞生的?时任中国科大化学物理系执行主任、现任中国科大合肥微尺度物质科学国家研究中心主任罗毅说:“这是天时、地利、人和的结果。”

科技自立自强之路

从“试错”到“智能创造” 机器人化学家来了

■本报记者 王敏

2 地利:多学科交叉的优势

创造出这样一位能阅读、能思考、能实践、能最终解决问题的机器人化学家,并非易事。

江俊认为,这主要得益于中国科学院集成攻关的传统和中国科大宽松的科研环境、多学科交叉的优势。

他回忆,大数据模型建立后,需要在实验中进行进一步验证,这意味着更大的投入。

2021年初,中国科学院发布了当年基础前沿和关键核心技术中的重大科学问题清单,“数据驱动的化学、材料和生物科学的机器人化学家”是46个项目之一。

江俊鼓起勇气,决定“揭榜挂帅”。很快,他的项目申请获得通过。从2021年到2025年,中国科学院将拿出2000万元经费,给予江俊团队连续5年的稳定支持。

也是在2021年,江俊向中国科大提交了一份“双一流”学科重点项目建议书。学校研究后,给予1000万元经费支持。同时,中国科大合肥微尺度物质科学国家研究中心、化学与材料科学学院出手相援,为研发团队提

供了研发场地、调配了研究设备。

“这种快速决策、快速支持、快速拨款的响应机制给了我很大激励,也为研究提供了及时、充分的保障。在人工智能领域,技术迭代速度极快,一步落后就会步步落后。”江俊说。

中国科大还有一个天然优势:这里会聚了不同学科背景的科研人员,他们可以优势互补、协同攻关,更加快速地把构想变成现实。

例如,机器人是整个平台中的重要一环。当初,尚伟伟与江俊商量后,愿意改变团队的研究方向,专门研发实验机器人。

安德鲁·库珀在参观机器人化学家实验室时,最羡慕的就是江俊他们拥有自己的机器人研发团队。江俊颇为自豪地说:“化学机器人指令集的每一行代码都是我们团队自己写出来的,如果要做一个新的动作,可以随时调整代码。安德鲁·库珀则必须请公司改代码,因为他的机器人是买来的。”

1 天时:科研发展的必然

2011年底,江俊决定回国加入中国科大,与他在瑞典皇家理工学院攻读博士学位的导师罗毅再次携手。

两人都从事理论和计算化学方面的研究,但各有侧重。罗毅主攻解析高分辨率分子化学成像,江俊则更关注材料设计和微观机理的研究。

2013年底的一天,罗毅与江俊一起出差。在路上,罗毅抛出一个问题:基于独立事例的理论计算,虽然符合传统研究范式,但效率太低。有没有可能通过大数据技术提高效率?

江俊经过调研后发现,我国严重缺乏科学数据的积累,在学术上非常被动。

“我们做了大量研究工作,但最终成果发表在外文期刊上。引用这些论文数据时还得花钱买,并且只有访问权,不允许大量下载。”江俊认为,要用好大数据技术,必须首先建立数据库系统。

从2014年起,江俊带领学生收集教科书、论文、专利中“沉淀”的化学经验和知识,用3年搭建起“中国人自己的化学材料大数据库”。

有了足够的数据库后,接下来的问题就是如何使用。在技术上,人工智能是处理大数据最好的手段。然而,这对罗毅、江俊而言是一个全新的工具。

不懂就学,知难而上。他们以自己熟悉的谱学、催化为抓手,应用人工智能方法开展研究,取得了良好的效果。值得一提的是,江俊发现的多个功能强大的描述符,被《科学》杂志两次全文推荐。

虽然融合了人工智能技术,但如何用大数据解决实际化学问题,依旧是个挑战。“因为数据本身缺乏知识逻辑的关联。”罗毅说。

2017年的一天,罗毅想到一个好点子:可否借助本校两位教授在发光材料、催化剂领域的深厚积累,将他们的知识与大数据库融合?

他激动地将江俊叫到办公室,说:“你把张国庆、熊宇杰发表的论文全部读一遍,这样就能形成一个‘张国庆大脑’、一个‘熊宇杰大脑’。研究发光材料时,就调用‘张国庆大脑’;设计催化剂时,就调用‘熊宇杰大脑’。这样一来,我们这些化学知识不足的人也可以用好数据了。”

就这样,江俊和同学们一起对海量文本数据进行人工标注,逐步赋予计算机自主阅读并分析文献资料的能力,仅用两年就成功培育出一颗“化学大脑”。

“化学大脑”有了,下一步就是想办法给大脑配上可以实际操作的“双手”。

2020年,英国利物浦大学安德鲁·库珀团队建造了世界首个移动机器人实验员。研读完他们发表的相关文章,江俊信心满满地对罗毅说:“我们不但可以让机器人做实验,还可以给机器人配上‘化学大脑’。”

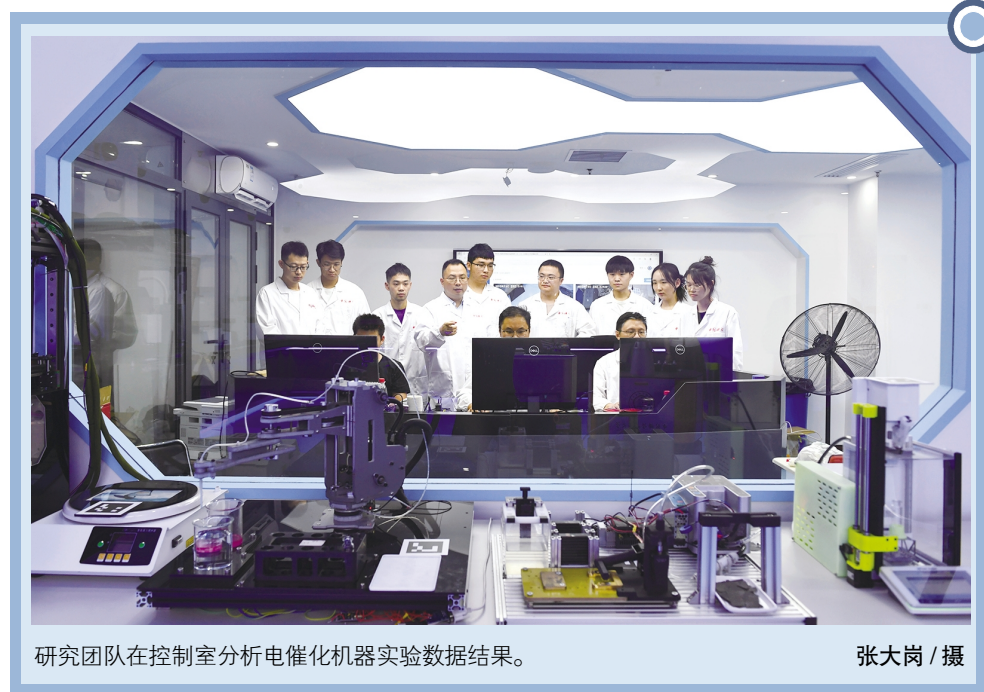
同年,在中国科大的一次交叉学科论坛上,江俊认识了该校信息科学技术学院从事机器人研究的副教授尚伟伟,两人在联手研发机器人化学家上一拍即合。

于是,在江俊的带领下,他们迅速组织了一个青年团队,夜以继日协同攻关。

2021年12月,全球首个数据智能驱动的机器人化学家,终于在中国诞生了。它融合“化学大脑”,集成了两台移动机器人、19个智能化学工作站和高通量计算系统,覆盖科学研究方法论全流程。

2022年9月,相关研究成果发表在《国家科学评论》上。审稿专家如此评价:这款机器人化学家的“机器人系统、工作站和智能化学大脑都是最先进的”,“将对化学科学产生巨大影响”。

“机器人化学家不是某一天一拍脑袋就能做出来的。”罗毅说,“这是我们与俱进、吸纳新知识、利用新技术解决问题的过程,也是持续进行科研探索的必然结果。”



研究团队在控制室分析电催化机器人实验数据结果。

张大岗/摄

4 表现:科研能力超强

数千年来,人类在化学领域不断解锁物质世界的奥秘,制造出许多令人惊叹的物品和材料。

然而,化学又是一门需要一点“运气”的学科。能产生反应的物质有很多,不同的量会带来不同的结果。更关键的是,反应釜完全是个“黑箱”,里面会发生什么、过程如何,在很多情况下是无法计算和计量的。

因此,在化学领域总有一些“无心插柳”的成果,也有一些“徒劳无功”的无奈。得到想要的结果,可能需要很长的时间。

机器人化学家的诞生有望改变这一状况。

江俊介绍,机器人化学家通过机器人精准的自动化操作能力,可以“不知疲倦”地执行任务。与此同时,“化学大脑”同步进行量子化学仿真模拟,再融合理论大数据和实验小数据产生具有预测能力的人工智能模型,最终得到算法预测、验证全局最优的实验配方。

正是这种理实交融的研究范式,能够从数百万种可能的配方中迅速识别最佳组合,极大加速了新物质的发现过程。

江俊以创制“高熵非贵金属产氧催化剂”为例,展示机器人化学家的超强能力。

高熵材料具有高混乱、高无序、高复杂度的特点,能提高能源电池的稳定性,对新能源发展非常重要。如果按照以往的方法,科研人员要从29种非贵金属元素中选出5种进行超过55万种配比组合,“试错”研究可能需要

1400年。

而机器人化学家通过阅读1.6万篇催化论文,自主遴选出5种非贵金属元素,并融合2.5万组理论计算数据和207组全流程机器人实验数据,建立并优化预测模型,将创制周期缩短为5个星期。

体现机器人化学家强大科研能力的,还有一项基于火星陨石的催化剂研究。

一直以来,火星移民是人类的梦想。实现这一梦想,首先要解决火星大气中氧气缺乏的难题。

就在前不久,机器人化学家利用火星陨石成功创制出实用的产氧电催化剂。该成果在《自然-合成》发表后,引发了很大反响。《自然》网站相关文章认为,此项研究“为火星探测和地外文明探索提供了新的技术手段”。

江俊介绍,如果用人工方式做实验,以5种不同的火星陨石作为原料,有超过376万个配方的排列组合。按每个实验验证至少5小时计算,找到最佳配方可能需要2000年。而机器人化学家只用了5个星期就做完了实验。给出的优化后催化剂配方,可以提供足够的氧气。

江俊等人亲切地把机器人化学家命名为“小来”,意味着开启未来物质科学探索的新纪元。它与人类智慧的结合将创造物质科学全新的可能,引领人类在星辰大海中不断拓宽知识的边界。

3 人和:年轻团队的助力

和机器人化学家一样年轻的,是研发团队——其成员以朝气蓬勃的“90后”“95后”为主。

“这群年轻人做出的东西比我想象的好得多,每一步的研究结果都超出预期。”罗毅说。

例如,在早期的数据采集工作中,还是中国科大少年班学院本科生的肖恒宇、霍姚远、冯超就表现出超强的研发能力。

刚开始,罗毅、江俊定期开讨论会,为学生们提供指导。但在讨论会上,他俩发现很多时候根本插不上话。之后经常出现的情况是,学生热火朝天地讨论,老师在一旁安静地听、欣慰地看着。

“不得不承认,信息科学时代,年轻人比我们的感觉更敏锐、理解力更强,他们的思维方式早就‘数字化’了。”于是,罗毅与江俊达成一致——把最重要、最困难的事情交给年轻人去做。

结果证明,他们的决定是对的。

在机器人化学家创建中,除了搭建数据库外,肖恒宇还自学计算机软件,开发出操作系统1.0版本。该操作系统既符合化学实验的操作习惯,又能发挥机器人的自动化优势,还能嵌入智能模型。

近期,肖恒宇又将操作系统升级到2.0版本,实现了硬件即插即用,模型普适迁移和操作云端共享。

专业工程师看了该架构都惊叹不已:“类似的框架只在西门子智能实验室看到过,只有经验丰富的工程师团队花费很高成本才能开发出来。”

人工智能领域有一句话:有多少智

能,背后就有多少人工。

培育“化学大脑”最重要的是赋予计算机理解物质的能力。当时,正在读硕士研究生的张百成,没日没夜地标注了几个星期,文献读得头昏脑涨,但完成的工作量微乎其微。

后来他改变“打法”,找了多位有专业背景的同学联合作战。他们一起构建化学词典,整理反应数据,很快设计出第一代文献阅读系统,得到了第一批高质量的标记数据。

“95后”赵路远结合本科学的电子工程知识,主攻化学工作站改造。她记得,与师弟曹嘉祺一起编写固体进样仪器的驱动程序时,由于信息掌握不充分,无法准确理解并调用仪器接口。他们多次咨询国外厂商和代理商,得到的回应都是“无可奉告”“商业机密”。

“后来,我们花了3个星期,重复做了3000次发送信号、等待反馈和记录响应实验,逐步逆向求出底层控制逻辑,成功实现了对机器人的远程控制。”赵路远说,国外厂商反过来主动联系他们,想购买他们研发的驱动程序。

类似这样的事,在机器人化学家项目攻关中并不鲜见。

令江俊欣慰的是,这群年轻人从不以专业限制自己,而是主动学习、迎难而上,想方设法解决问题。

眼下,这群学生已成为实验室的骨干,各有各的绝活儿。其中,肖恒宇正在设计集中式大规模机器人化学家平台,张百成在赋予机器人谱学智能,赵路远在探索适配产业应用的化学工作站。

5 目标:数据与智能驱动的化学研究新范式

放眼全球,在机器人化学家领域,目前只有英国格拉斯哥大学和中国科大实现了自动化文献阅读功能,英国利物浦大学和中国科大实现了机器人可移动性。最为重要的是,国外同行研发的同类产品均未实现文献调研、合成、表征、测试、数据处理全流程,可进行的研究类型较为单一,离智能化尚有差距。

“这说明我国的机器人化学家研发与国际同行处于并跑阶段,并在某些方面具有领跑优势,发展潜力巨大。”江俊说。

从机器人化学家的成功经验看,罗毅认为,未来可以将科学工程化,把优秀科研人员聚集在一起,以机器为载体,以数据为要素,做有组织的科研。

罗毅表示,他们的目标是迅速扩大机器人化学家这一技术平台的规模,增强其能力,建成世界上最好的“化学医院”。

“我们希望可以扩大平台规模、多地联动,建成不同形态的智能机器人,

比如天上飞的、地上跑的、桌面微操作的机器人,日均完成万次以上化学实验操作,满足航空航天、医疗卫生、材料科学、化工原料、生物医药等领域的研发需求。”

目前,研发团队正在赋予机器人对化学过程的感知、预判能力。“比如,赋予机器人谱学智能。以前是训练机器人如何做实验,现在希望机器人通过谱学智能直接看到微观世界的变化,这样就能够更快地预判化学变化,超越人对实验的理解。”罗毅说。

罗毅期待更多科研人员将重要的科学问题放到机器人化学家这个平台上,以提高科研效率,产出更多新成果。

“我们处在一个技术快速迭代的时代,要主动拥抱新技术,以积极开放的心态迎接世界的变化。”罗毅这样阐述团队的终极目标:创造出具备创造力的智能化学家,及早形成数据与智能驱动的化学研究新范式。



江俊在机器人化学家实验室里调试操作指令。

张大岗/摄