

建立规则，让AI进化风险可控

□ 郑念

正念思维

近日，未来生命研究所(Future of Life Institute)公布一封公开信，呼吁所有AI实验室立即暂停训练比GPT-4更强大的AI系统，时间至少为6个月。信中质疑，人类是否要冒着失去对文明的控制的风险，使人工智能系统最终在数量、智慧和技能上超越、淘汰和取代人类？AI开发者必须与政策制定者合作，大力加强发展健全的AI治理体系。近期火爆全球的ChatGPT拥有语言理解

和文本生成能力，会通过连接大量的语料库来训练模型，在交流时与真正人类几乎无异。ChatGPT不单是聊天机器人，还可以写作科研论文、处理和进化自然语言、制作视频、图片、编写场景故事等。这意味着，ChatGPT把人类进化模式尤其是知识生产的模式推进到了一个新时代，可以想象，未来的知识产出也将开启爆炸式增长模式。这将对科研、教育、生物医药、科技情报等领域产生重要影响。

目前，ChatGPT进化速度很快，能力和智力提升速度也很快。据测试，GPT-3.5的智能与中等智力的人相仿，但在短短几个月

后，GPT-4可在考试中超过九成人类。进一步讲，人工智能+基因工程或将导致“超人类”的诞生。未来，人类将在智力、体力、生存能力等方面面临巨大的挑战。

有人预言，照此速度进化，人工智能的奇点将会很快到来，其技术文明将对制度文明带来严峻的挑战。如何平衡技术文明和制度文明的关系这一话题，受到广泛关注。其实，ChatGPT已经不是单纯的技术问题，它深深地刺激和拷问着人类的灵魂和伦理道德。这再次说明，技术需要与人文相结合，才能在正确的轨道上发展，人文是科学发展的导向，需要正念思维的引导和规制。如果

在伦理上不被严肃对待，缺乏约束性的进化，ChatGPT可能成为人类为自己生产的难以制约的敌人。这就要求我们在思维层面发挥优势，以避免人机竞争的结果发生，抑制资本因单纯逐利而选择智能机器，科技因单纯的追求技术而毁灭人类的未来。

总之，新的挑战将是全方位的，我们必须有清醒的认识。无论是多么智能的技术，其在进化的过程中总会带来风险。技术越进步，智能化的水平越高，就越需要正念思维的指导。正如公开信所言，“让我们享受一个漫长的‘AI夏天’，而不是毫无准备地进入‘秋天’”。



左起：黎介寿、黎黎、黎磊石 (图片由作者提供)

清明将至，祭奠逝者，追思故人，我不禁想到了今年1月去世的中国工程院院士、国际著名普通外科专家、医学教育家黎介寿。

说起黎介寿，不得不提“一门三院士”的“黎氏”传奇。上世纪初，黎介寿和他的哥哥黎黎、弟弟黎磊石三兄弟出生在风雨如晦、备受外国列强凌辱的旧中国，后来成长为日益强大的新中国。伴随着祖国的发展与进步，他们在医学领域都作出了开创性的贡献，先后当选为中国工程院院士，成就了中国医学界“黎氏三院士，悬壶济苍生”的一段传奇和佳话。

黎介寿是中国烧伤医学开拓者，在中国烧伤学素有“百科全书”的美誉。

1958年，“大炼钢铁”时期，我国烧伤病人陡增。黎黎意识到，烧伤是威胁民众生命健康的大患，他主动承担重任，提交了一份烧伤防治研究的请战书。当时的第三军医大学当即决定由他牵头在普外科组建烧伤救治队。

1966年，四川合江一支钻井队不幸遭遇井喷，数十人被烧得面目全非。黎黎率领一支精英救治小组赶赴现场投入到抢救战斗中。他本人更是几天几夜没合眼，将一个个烧伤患者从死亡线上夺了回来。

1994年，黎黎被美国烧伤协会授予“伊文斯”奖，并当选中国工程院首批院士。他创建全军第一个烧伤中心和第一个烧伤研究所。

黎介寿，1924年10月出生，1963年4月入伍，1996年当选为中国工程院院士。

黎介寿一生大爱，一生攀登，一生厚德，一生为梯。在人类肠道疾病的“神秘王国”，他穷尽一生心血，不懈攀登科研高峰，成为“世界上研究肠子时间最长的人”。他日以继夜、相依为命的坚守，换来小肠移植这个世界性难题的攻克。在“九曲回肠”的世界里，黎介寿创造了一个又一个“亚洲第一”“世界第一”，成为我国肠外瘘治疗的鼻祖、临床营养支持的奠基人、亚洲人同种异体小肠移植的开拓者、加速康复外科和损伤控制性外科等技术理念的倡导者。

黎介寿从医70年，用精湛医术、高尚医德和仁爱之心，为患者消除痛苦、送去健康。“只要病人还有一口气，做医生的就不能轻易放弃。”黎介寿视拯救生命高于一切，只要有1%的希望，他就尽100%的努力。他先后做手术上万台，收到患者和亲属赠送的表扬信和锦旗14000多件。他先后带教了350余名硕士、博士，后都成长为我国普通外科领域的佼佼者，为党的医学事业集聚了一批优秀人才。他对学生说：“现在的年轻人有很多天时地利的条件，生活在黄金时代，如果不努力，对不起国家，对不起人民。”

2014年，浩瀚星空又闪耀一颗小行星，他的名字叫“黎介寿星”。这颗小行星位于狮子座，距离地球约2.21亿公里。

黎介寿，1926年10月出生，曾任南京军区总医院副院长、全军肾脏病研究所所长，1994年当选为首批中国科学院院士，国际著名肾脏病学家，是中华肾脏病学会及亚太地区肾脏病学会创始人之一。他是首批国家政府特殊津贴获得者，先后获得军队专业技术重大贡献奖、“何梁何利”基金奖、中国医师奖，荣立一等功1次、二等功6次、三等功8次，获国家科技进步奖6项，军队科技进步一、二等奖12项。他领导的肾脏科是军队、江苏省医学重点之重的学科及实验室。2010年3月16日，黎磊石因癌症长期医治无效，在南京逝世，享年84岁。

黎黎、黎介寿、黎磊石兄弟3人分别担任烧伤、普通外科、肾脏病3个全军专业组的组长以及3个全军研究所的所长，都在各自领域取得了举世瞩目的成就。

三兄弟曾经有一个共同的约定：生，共饮长江水；死，共聚长江魂。黎黎、黎磊石过世后，遵照他们的遗嘱，骨灰撒于滚滚长江。

如今，中国医学界的“黎氏”传奇在滔滔江水中重聚……

(作者系国家教育咨询委员会委员，中国科技馆原馆长、研究员)

黎门三院士 悬壶济苍生

□ 王渝生

大树叶茂源于根深

——对科普创作中若干关系的思索(下)

□ 卞毓麟

科普对象不同，特色势必各异。少儿科普与成人科普之特征有何异同，就值得推敲。科普层次众多，分工必有侧重。一线科学家打造的“元科普”作品，乃是科普这棵大树的根。

少儿科普与成人科普传播规律不同

少年儿童是未来的主人。“少年强则国强”，未来的将军、科学家、艺术家都是今天的孩子们。因此，对少年儿童的科普教育确实极为重要。那么，成年人的科普又如何呢？

我无意在“成人科普”与“少儿科普”之间画一道生硬的分界线，而是想传递这样的理念：应该看到，主导今天、影响未来的乃是现在的成年人。因此，对于成人科普也不能掉以轻心。不过，人们通常并未把“成人科普”作为一个整体性概念来考察，而是比较习惯于强调其中的这一部分或那一方面。

近30年来，我们已逐渐熟悉国际上流行的“公众理解科学”这一概念。从对公众科学素养调查的对象(18岁以上的成年人)和内容(包括对科学技术知识的了解程度、对科学技术的态度、对国家科技政策的看法，获得科学技术信息的渠道和手段等)可以清楚地看出，它大致就对应于成人科普。

值得注意的是，成人科普和少儿科普具有不同的特征，应该遵循不同的传播规律。少儿科普比较侧重于认知性。比如，低龄儿童看识字字、数数算算，中学生阅读的科普书。同时，少儿科普也较侧重于知识的基础性，少年儿童今后的路还长，要一步一步一个脚印地往前走，对他们的科普应该从比较基础的东西开始。



3部广受好评的元科普佳作(左起)：汪先著《深海浅说》，吴家睿著《生物学是什么》，冯伟民著《进化史诗16讲》。(图片由作者提供)

相对而言，成人科普更多关注科学人文。成人有社会阅历，要在社会舞台上进行更有效的活动，就更需要熟悉科学的人文方面。同时，由于当代科技进展日新月异、层出不穷，成人科普很难按部就班地往前走，因此科普内容更应注重时效性。

不管是成人还是儿童，当然都喜欢有趣的东西。如果把儿童喜爱的有趣称为“好玩”的话，那么针对成人的有趣实质上更倾向于“易受”。成人科普的趣味性，在很大程度上体现为它是否容易被接受。

“元科普”奠定坚实基础

什么是“元科普”？它是工作在某个科研领域第一线的领军人物(或团队)生产的科普作品，这种作品是科学家对本领域科学前沿的清晰阐释，对知识由来的系统梳理，对该领域未来发展的理性展望，以

及科学家亲身沉浸其中的独特感悟。

为什么用“元”这个字呢？查《辞海》即可知，“元”的主要意思包括：“始、第一”；“为首的”；“本来、原先”；“主要、根本”等。例如，元旦、元年、元首、元勋、元老等等。用在科普上，“元科普”就是科普中的元典之作。

科普好比一棵大树，“元科普”就是这棵大树的根。或者说，它是源自科学前沿团队的一股“科学之泉”。

“元科普”为其他形形色色的科普作品提供坚实的依据，包括可靠的素材和令人信服的推理，真实地传递探索和创造过程中深深蕴涵的科学精神。

例如，中国科学院院士、著名海洋地质学家汪先著《深海浅说》一书就是个典型。汪先生82岁高龄时还亲自深入南海，发现深水珊瑚林。《深海浅说》从深海

的基础知识开始，一直讲到深海的开发利用，既反映最新进展，又揭示科学发现的过程，填补了中国深海科学的“元科普”空白。

近年来，我国的元科普佳作接连问世，令人欣喜。例如，吴家睿著《生物学是什么》、冯伟民著《进化史诗16讲》等，亦皆为精品。或有人问：一线科学家都是很忙的，他们又如何来做科普呢？这当然因人而异，但也有共同点，那就是应该优先把有限的可支配时间与精力用来做别人难以替代的、因而潜在社会影响也最大的科普，这就是“元科普”。

早在1995年，我就在《“科学宣传”六议》一文中谈到，科学技术的前沿知识和最新发展，首先只能由科学家来传播。“在整个科学传播链中，科学家是无可替代的‘第一发球手’。”“当然，有了‘发球手’还要有‘二传手’，有了‘主角’还要有‘配角’。这样才能调动社会各方面的积极性，唱好科学宣传这台戏。”

发球很重要，二传同样很重要。有了这些“二传手”，才能把科学之球传到全社会的方方面面，深入千家万户。

卓越的科学家竺可桢也是一位杰出的科普家。他以身作则，号召大家“做一个光荣的科学宣传员”。今天，无论是我们的科学家还是科普人，都应当以竺可桢为榜样，把“做一个光荣的科学宣传员”作为自己始终不渝的崇高追求。

(作者系中国科普作家协会前副理事长)



室温超导离我们还有多远

□ 陈思进

近日，罗彻斯特大学的物理学家兰加·迪亚斯在美国物理学会会议上宣布，在近环境压强下实现了室温超导。随后，我国南京大学超导物理和材料研究中心提交一篇新论文，证实了迪亚斯的实验结果。2022年，该室温超导新突破被证实后，室温超导离我们究竟还有多远？

不同元素可在不同温度下实现超导

超导是指材料在低于某个温度时电阻变为零的现象。如果某种超导材料通过电流而不会有电阻的电流传输，那么这“某个温度”就被称为超导转变温度。

出现超导时一般会存在两个现象，一是电阻为零；二是完全的抗磁性。也就是说，当材料在一般状态下，金属磁感线可以穿过它，但是一旦进入超导状态，磁感线就被排挡在外，此时就会产生一股巨大的甚至是它自身重量几百倍的排斥力，足以将该材料推起来。磁悬浮就是利用了这一原理。

1911年，荷兰物理学家海克·卡末林·昂内斯在实验室中发现超导现象，并因此获得

了1913年的诺贝尔物理学奖。昂内斯是在零下269摄氏度(接近绝对零度)的温度下发现了水银超导现象。

之后，科学家逐渐发现，目前的118种元素，通过不同的元素搭配，能在不同的温度下实现超导。经过多年的努力，超导的温度从接近绝对零度上升到零下200摄氏度、零下150摄氏度等。

科学家还发现，超导现象可在各种不同的材料上发生，包括单纯的元素如锡和铅，以及各种金属合金。但是，超导现象不会发生在贵金属如金和银上，也不会发生在大部分的磁性金属材料上。

超导在医学、工程等领域广泛应用

由于超导材料的电阻为零，这意味着在超导材料中通过电流不会损失能量，因此可以连续流动而不需要输入任何外部能量。基于该特性，超导材料在电力传输和磁共振成像(MRI)等领域中有着广泛的应用。

超导现象的发现给物理学带来了重大的贡献，也推动了科学技术的发展。例如，超导技

术在MRI中得到广泛应用。MRI利用了超导材料在外部磁场下的特殊性质，可以在不使用有害的放射性物质的情况下生成人体内部的高分辨率图像。

超导是一种令人兴奋的现象，它在医学和工程等领域中具有极其重要的应用，甚至将会为我们的生活和工作带来革命性的变化。

然而，超导材料的应用仍然受到一些条件限制，如需要在极低的温度下才能实现超导等等。近年来，科学家们对高温超导材料进行了研究，这些材料可以在相对较高的温度下实现超导。这种研究为超导材料的应用提供了新的机会和挑战，也推动了超导技术的进一步发展。

室温超导将带来能源自由

室温超导给我们带来无限遐想。以磁悬浮列车为例，如果能实现室温超导，目前使用的汽油、柴油甚至新能源汽车等交通工具将面临淘汰，人们只需要在地下埋上磁感线圈，车辆便可以用上磁悬浮，而且这些

交通工具可使用直流电来代替交流电。一旦实现室温超导，将会给人类社会带来什么变化？

首先，超导材料可以在电力传输中减少能量的损失。如果实现室温超导，我们将能够制造更为高效的电力输送系统，这将减少能源浪费和环境污染。

其次，在当前的电子设备中，电阻会导致能量浪费和热量产生，从而限制其性能。如果实现室温超导，我们可以开发更高效、更快速和更强大的电子设备。

第三，由于超导材料可以在磁场中产生极强的力，如果实现室温超导，我们将能够建造更快速、更安全和更环保的高速列车和飞行器。

室温超导一旦实现，将会推动人类进入到能源自由的新时代。

(作者系科幻作家)



当脑中印象复原为现实中图像

□ 尹传红

信息和实施监控的目标。“神经资本时代”到来。当新技术可以便捷地侵入并控制人的大脑，将人的思想和意识连接起来，那么，控制我们行为、个性的大脑区域，岂不与计算机及数据存储器也连接起来了？

整整100年前，1923年，在英国作家埃德温·奥德创作的小说《发条人》中，一个由螺栓固定的机械装置(时钟发条)被植入到一个人的大脑里，全方位地操纵着这个人的行为，但同时也赋予了他进入其他维度空间旅行的能力。此中所采用的技术手段，未免简单粗暴，作者意在批判身体的机械化及世人对技术的盲从和滥用。

1960年，美国科学家曼弗雷德·克莱因斯和内森·克莱因首次提出了“生化人”(Cyborg)的概念，设想未来会出现一个能够自我调节的“人机系统”，其中人类部分由药物或者调控装置控制机体，使其在离开正常环境后仍然可以生活，譬如在漫长的星际旅行中对外太空的适应。后来，Cyborg(又译赛博格)这个词被专门用于指代机械化或电子化的人类，它实际上是英文中神经控制装置

(cybernetic device)与有机体(organism)的混写。

自21世纪初以来进行的诸多实验中，一种叫作脑门(BrainGate)的设备被植入到几个不同类型的偏瘫患者身体中。此装置将大脑直接与一台电脑相连，把思维以脑波的形式传到电脑，电脑再将脑波转换为动作。这样，患者便能将脑波转换为个人想要进行的动作。

随着脑的各种机制正在被加速阐明，脑各个区域的功能和作用得到了明确，利用人工智能(AI)技术解码神经网络交换的神经信号渐次成为可能。脑解码(brain decoding)技术已经可以解读大脑中所想的图像、动作以及梦，人之脑海中的画面也己能够再现，即把脑中的印象复原为现实中的图像。换句话说，通过脑解码技术已能“读取”一部分人的心中所想，并将其再现在计算机屏幕上。

日本京都大学教授神谷之康在2008年曾做过这样一项实验：先让受试者观看400张点阵图像，并让AI学习当时的功能性磁共振成像(fMRI)影像数据。当受试者看到新的点阵图像后，将其大脑的fMRI影像输入

AI，AI便能利用事先学习发现的规律再现点阵图像。尽管再现之成像细节部分的形状和颜色有所不同，但从整体上看，与原来的图像颇为相似。

随后，神谷又尝试在不给受试者看任何图像的情况下，再现其脑内想象的画面，也获得了初步成功。再进一步，利用同样的原理尝试对一部分梦境进行解读，将睡眠中的脑活动成像。神谷2013年在《自然》杂志上发表论文报告说：AI能够以较高的准确率预测特定物体是否会出现受试者的梦中。

对大脑实施操控或自我优化，医学上的利好不得不言，但同时也会引发争议。早在2012年举办的一次信息安全大会上，脑电波中的安全隐患就被提了出来。英国、美国和瑞士的科学家联合发布的一项研究成果，展示了恶意软件如何通过目前已有的商用脑机接口设备获得用户的个人信息。他们设置的脑电波信号探测模型，在开始了一系列“黑客测试”后，几乎能将用户的密码、开户银行以及位置信息等一一泄露。

脑机融合，不只是带给我们美妙。那是一把双刃剑。



在大脑深处植入芯片，获得更强的脑力，解决诸多问题……这些科幻小说中的“设计”，已然有了现实的雏形。连接大脑和机器的脑机接口技术，眼下正加速走向实用化。

继埃隆·马斯克旗下的“神经连接”公司之后，一家名为“同步”的生物技术在脑机接口领域也迅猛崛起。2022年，该公司研究人员成功地在一名肌萎缩侧索硬化(渐冻症)患者的大脑中植入了一种运动神经假体，并使其得以通过监视器进行通信。在新近的最后轮融资中，它又获得了7500万美元的投资。

看起来这场脑机竞赛才刚刚开始，西方社会已经表达出这样的忧虑：两者开发的脑机接口技术，都在朝着远程干预我们的大脑以及我们的思维发展，最终实现提取敏感信息



余生趣谭

三人行必有我师 马国馨院士篆刻