

2016年发生的两件事，让心理学家刘嘉奋斗了20年的学术信仰崩塌了。

第一件事是世界顶尖围棋高手李世石以1比4败于人工智能(AI)棋手AlphaGo。这让刘嘉猛然发现，“一个全新世界正缓缓拉开帷幕。”

第二件事则是刘嘉的亲身经历。在他参与策划与制作的江苏卫视《最强大脑》节目中，面对通过幼年人脸识别20年后的成年人、双胞胎识别、现场实时照片识别等挑战，人类顶尖选手败给了AI。

他本来自信满满。早在2002年，他在美国麻省理工学院(MIT)进行博士学位答辩的题目是《面孔识别的认知神经机制》，并在第二张PPT里写道：“现在最先进的机器识别面孔的正确率只能是随机水平，而人类能够在几秒钟识别上百张面孔。为什么人类如此伟大？为什么人类如此聪明？为什么机器如此愚笨？”

在策划《最强大脑》节目之初，他的判断是，人类肯定更厉害，可以不费吹灰之力碾压AI，这次要“让人类的智慧在镜头前痛痛快快地羞辱一下AI”。

当AI战胜人类的那一刻，他在沮丧之余顿悟：他一直追求的“面孔识别问题并非人类智能的本质所在，自己围绕一个假问题奋斗了20年！”

他决定躬身入局，去寻求并攻克真正的问题。为了新的目标，刘嘉辞去了北京师范大学心理学部部长的职位。他要重返实验室，做一个全职科研的普通教授。

新的信仰

时间拨回到1997年。刘嘉刚到MIT脑与认知科学系开始攻读博士学位，他和导师南希·坎韦施有一则这样的对话。他说，自己想挑战大的课题。南希告诉他，视觉是人类所有的感觉系统中最主要的信息来源，有60%的大脑区域与视觉有关，而在所有的物体识别中，面孔识别是最难的。

刘嘉下定决心：既然面孔识别是那颗星冠上的明珠，他要做那个摘取明珠的人。从此，5年读博生涯以及回国之后的15年间，他一直深耕面孔识别。他想知道人是怎么识别面孔的，大脑的什么区域在执行这个功能，以及执行的过程。

在读博期间，他发现了重要的结果：当人们看见一张面孔，大脑会用100毫秒把面孔和非面孔区分开，然后才会把这个信息传输到一个专门用于面孔加工的区域，之后再再用70毫秒把这个面孔识别出来。

这就是面孔识别的独特之处，跟其它物体的识别是完全不同的。这个功能无比重要——只有在人群中迅速找到母亲的面孔，婴儿才能提高自己的生存概率。这篇研究文章发表在《自然-神经科学》，迄今引用接近900次。

然而，他一直以为的最本质的智能问题竟然被AI攻克了。更何况，2016年的AI在人们看来还特别弱智。显而易见，他选错了方向。

他追求的是终极问题，面孔识别并不是。那问题是什么呢？当然就是刻在阿波罗神庙的那句“认识你自己”：我们是怎么理解自己的，我们是怎么理解这个世界的，它背后的第一性原理是什么？

44岁的刘嘉开始寻找新的线索：在“人猿相揖别”的数百万年里，人类主要进化以语言为载体的思维能力，所以语言和思维才是最为重要的。

2022年11月30日，大语言模型ChatGPT面世的那天，刘嘉正发着烧，有朋友给他发了一个GPT-3.5的链接。最初的问题他已经不记得了，但他立即意识到这大概就是通用AI(AGI)的雏形。震惊之余，他的病都好了一半。

再延伸出去，他得出了一个更为震撼的结论：那一天是AI进化史的分界线。在此之前，AI只是一个工具；在此之后，地球上有了另一个能理解人类的“物种”，人类不再是孤儿了。

9年前，读博的刘文丰在查阅文献时萌生了一个想法——将不同维度的水资源要素整合起来开展研究。当时这个设想有些“超前”，因缺乏系统数据与分析框架难以推进。但刘文丰没有放弃，而是把它深埋在心底。

如今，这颗种子终于破土而出。中国农业大学水利与土木工程学院教授刘文丰所在的研究团队首次提出了考虑“蓝水-绿水-水质”的多维水短缺概念及其评价框架，推动了水短缺的评价体系从“单一维度”到“多维度”评价体系的转变，为未来精准识别、分析与应对水资源危机奠定了理论基础并提供了实践路径。近日，相关研究成果发表于美国《国家科学院院刊》。

眼前一亮

在传统的水资源短缺研究中，科学家通常聚焦于三种水体——源自江河、湖泊和地下水等可以被直接取用的“蓝水”，储存于土壤中供植物吸收的“绿水”，及受污染影响而需要处理的“水质水”。三者相互关联，共同决定人类和生态系统的实际可用水状况。但在过去的水资源研究中，这三者往往被割裂评估。

2016年，在瑞士苏黎世联邦理工大学攻读博士学位时，刘文丰在一次查阅文献时，读到了一篇荷兰瓦赫宁根大学研究者的论文，其首次尝试将蓝水与水质耦合分析。尽管篇幅不长，模型也不复杂，却让他眼前一亮。“它触及了一个一直在我脑中盘旋的问题——如果在蓝水-绿水研究的基础上引入水质维度，不就能建立一个更完整的水资源短缺评估框架了吗？”他立刻和导师草拟了一份概念性方案，试图搭建一个涵盖蓝水、绿水与水质水质的综合分析框架。

然而，由于缺乏数据支撑，这个方案没能发表。如今，刘文丰将多年前的想法付诸实践。他和研究团队通过构建全球尺度的水文-



刘嘉

受访者供图

刘嘉的学术信仰之跃

■本报记者 孙滔

刘嘉说，“如果达尔文看见这个时刻，他一定会说，一个新的生命诞生了。”那么，如果脑科学加上人工智能，会不会在某一天造出一个数字人类大脑呢？

彼时，他的身份已经是清华大学基础科学讲席教授、北京智源人工智能研究院首席科学家。

他是如此激动，以至于为此写了一本名为《通用人工智能：认知、教育与生存方式的重构》(以下简称《通用人工智能》)的书，并于今年6月出版。在自序中，他写道：“把这近十年中我的摸索、挣扎、领悟和疑问，攒成了文字，与你分享。”

一面镜子

如果说ChatGPT的问世给了刘嘉新的信仰，那么2024年诺贝尔物理学奖颁发给机器学习则给了他极大的信心，因为作为获奖者之一的“深度学习之父”杰弗里·辛顿正是心理学出身。

不仅如此，辛顿的履历也让刘嘉有极大共鸣：英国剑桥大学入学后仅一个月，辛顿便厌烦了物理学“每天12小时实验与誊写笔记”，他愤而辍学。一年后重返学校，他开始学习生理学，又很快产生了不满，因其只涉及轴突电生理而不涉及大脑工作机制，于是转向哲学。1970年，他再度转向心理学并获得学士学位。剑桥大学毕业后，辛顿成了一名木匠。这期间，他读到了加拿大心理学家唐纳德·赫布的《行为的组织》一书，深受触动。他决定用仿生方法研究大脑智能，之后进入英国爱丁堡大学攻读博士学位。

这样的人生简直就是刘嘉的一面镜子，他也曾在人生追求上摇摆过。

刘嘉是重庆涪陵人，小时候经常在夏夜铺着凉席睡在室外，仰望星空的时候常常好奇宇宙的广阔面貌，这使得他在高中一直想读天体物理专业。不过他的想法因为一本弗洛伊德的传记改变了。他忽然觉得心理世界比宇宙更为广阔、更为复杂，他要去学心理学。

1990年进入北京大学心理学系后，他才发现这门学科太“文科”了，充满了哲学思辨以及没有逻辑的死记硬背，远非他想象中的科学。于是在本科期间学了北京大学计算机系的核心课程。他还辅修了无线电系的电子学与信息系统，皆因大脑也是一个信息加工的系统，他觉得学信息论或许有助于理解大脑。

他最终坚持完成了心理学的学业，并在读完北京大学认知心理学硕士后，奔赴MIT攻读博士学位。

其实，刘嘉在大学做的事情远不止上述。按他的说法，“红道”“黑道”和“黄道”他都试过，最终才回归心理学。

所谓“红道”，就是“从政”。他大一的时候竞选北京大学的校学生会主席，在第二轮即被淘汰出局。在这个过程中，他只感到累，完全找不到“high”(兴奋)的感觉，由此明白了这条路不是自己想要的。

“黄道”跟挣钱有关。在1995年硕士一年级的时，他就靠写软件挣到了不少钱。他不会理财，有了钱就会带朋友去北京大学南门“胡吃海塞”“喝了就吐”。他仍然没有“high起来”。

他决定坚守“黑道”，也就是做学问。这条路虽然很累，报酬也很低，但“越做越high”，这就足够了。

刘嘉最早接触AI是1994年。他用传统人工智能方法符号AI做了关于人格测量的专家系统，并获得了北京大学“五四青年科学奖”一等奖，他是唯一获奖的本科生。他那时看到的AI远不是他期待的，直到他听了一位刚从日本回国的博士讲授人工智能课程，才感受到了AI的“灵动”。可惜那时电脑算力太小，一台386台式机的算力远比不上现在的智能手机。

之所以奔赴MIT读博，其一是因为刘嘉通过互联网知道了人工智能之父马文·明斯基，后者当时在MIT人工智能实验室任教；其二是他读了已故计算神经科学领域先驱大卫·马尔关于

视觉的著作。在那本书中，这位MIT教授用了一堆数学公式讲述人类为什么能看见这个世界。刘嘉顿生崇拜，认为认知科学就应该有这样的，那才是自己应该为之奋斗的方向。

可惜的是，当刘嘉在MIT见到马文·明斯基，并聊起人工智能神经网络时，这位大佬却一脸消沉，告诉刘嘉这个方向不值得做，然后建议他回归脑科学研究。

搭乘冷冰冰的电梯从马文·明斯基办公室出来，他的心绪一如马文·明斯基般消沉。他还是听从了大佬的建议，再也没碰AI。

“理解智能的本质”

其实刘嘉对AI的兴趣一直没有消亡。以前的AI太弱智，没法做好玩的事情，当AI不再弱智了，做事的时机来了。

回顾最近十年的学术之路，回归实验室后的刘嘉并不顺利。他反思，自己犯了多个错误。

第一个错误是，为了追求可解释性，他使用的人神经网络的参数量太小。他没有关注美国OpenAI公司的工作，而是追随了谷歌DeepMind XLand的工作，甚至在2020年6月GPT-3问世之后，他曾批判那是个试图靠堆参数堆出智能的“傻大姐”。他坦承，自己错得格外离谱。

这个错误，他是在2023年才认识到的。毕竟当神经网络还太小时，不可能具有真正的智能，人们没办法指望从只有302个神经元的秀丽隐杆线虫中得到关于智能的研究结果。

第二个错误是跟风。他也想做基于Transformer的大模型，直到2023年底，他才想到了生物进化的维度。

他发现，更多的神经元数量并非高级动物区别于其他动物的唯一特征，否则神经元数量远多于人类的大象和鲸会比人类更聪明，那么进化就成了他参照的智能法则。他要回到自己擅长的领域里，把脑科学和AI结合起来，看看智能是如何进化而来。因为他的第一性原理是“进化不会干无聊的事情”。

与如今流行的智能科学(AI for Science)不同，刘嘉想要做的是，让脑科学帮助AI来发展，就像当年辛顿一心要让机器模仿大脑的运作机制一样。

刘嘉看到了当下大模型的短板，那就是，它只能做从1到100的事情，而对从0到1的颠覆式创新的事情无能为力。所以在他看来，如今最该突破的反倒是脑科学，只有把脑科学的根本问题解决，才能启发下一代AI的诞生。

刘嘉在做什么呢？“理解智能的本质”，刘嘉课题组在做什么呢？

在刘嘉的实验室，《中国科学报》记者见到他们正在利用宽场的单光子钙成像技术记录小鼠视觉皮层的上万个神经元的活动。刘嘉想知道，当小鼠在做某个认知任务的时候，这些神经元构成的生物神经网络是怎么活动的。这样的研究或许能启发新的人工智能神经网络算法。就在前两年，他们课题组还在以猕猴为研究对象。不过他很快发现，每次能记录的猴子神经元有限，于是他转向了小鼠。这样他就能一次观察上万个神经元的活动，只有如此，他才能把研究结果和人工智能神经网络作对比。在做神经元成像时，同时记录100个神经元和同时记录1万个神经元，得到的结果是完全不同的——就像大模型的发展一样，规模会带来认知功能的涌现。

“信仰之跃”

如今的刘嘉不出差，不参加各种会，“老老实实在第一线做研究”。

刘嘉这种坦然的姿态传承自他的导师南希——一位个性十足的科学家。当年回国之际，刘嘉问导师，建立实验室有什么诀窍。南希告诉他两个关键词——tough and fair。刘嘉不解：fair是公平的意思，那tough是对人严厉吗？南希说，

其实是simple and straightforward(简单和直接)的意思。

刘嘉马上明白了，导师的意思是，不需要转弯抹角，绕来绕去。若要去申请经费，拿文章、拿实验提案就足够了；若是学术观点和别人不同，就可以直接表达出来。fair则是对事不对人，科学面前，学生、老师平等。

但是，tough and fair并不意味着特立独行。他意识到，DeepMind战胜MIT与哈佛大学的布罗德研究所，因解析蛋白质结构拿到了2024年的诺贝尔化学奖，充分说明了传统小而美的实验室和传统老师、学生的单线合作研究模式已然失效。

在AI时代，科学必须得像DeepMind那样以工程化的方式来推动，即Engineer for AI for Science(基于工程架构的人工智能助力科学研究)。跨学科的合作，已经成为脑科学+AI发展的必要条件。

2024年，他开启了“小曼哈顿计划”，即在学校的支持以及已有心理学基础上，创建了融合心理学、脑科学和人工智能三个交叉学科的心理与认知科学系，开启了一段新的征程。

“信仰之跃”，这是刘嘉为辛顿写的小传的题目，其实那是写给他自己的。

辛顿是刘嘉最崇拜的偶像。当辛顿攻读博士学位之际，正是神经网络研究的严冬。马文·明斯基在其著作《感知机》中严厉批判了由神经网络通向AI的道路，断言“研究神经网络是一条死胡同”。这时候，辛顿的导师也背弃了神经网络，转投符号主义AI阵营。然而辛顿却执着于神经网络研究，成了远离经费、导师和学界支持的孤勇者。

辛顿的信仰是坚定的，“没有理由认为人工神经网络做不到我们能做的一切”。早在剑桥大学读书的时候，他就宣告：“理解大脑的唯一办法，就是亲手造一个。”这是他信仰。

当年刘嘉辞去北京师范大学心理学部部长的职位，这在很大程度上有相似之处。

他把这个信仰的根源归纳为延迟满足，这种延迟满足甚至是超越生命周期的。刘嘉经常告诉学生，做事情要考虑时间尺度，要把它放到5年、50年，甚至超越你的生命的尺度上来看待其价值。

在方法论上，他极力推崇的一个策略是，进攻是最好的防守。

他是国内使用功能磁共振研究大脑功能的首驱之一。当年他刚到MIT读书时，导师就开始用功能磁共振研究大脑功能。刘嘉于2002年加入中国科学院生物物理研究所，参与组建了国内第一个用于脑科学研究的磁共振设备。2006年到北京师范大学后，他主导建设了国内高校第一个磁共振脑成像中心。

他没有满足于这些成就，反而跳出了舒适区，再也不能磁共振了，原因是其空间和时间分辨率太差，完全不能满足人工神经网络研究需要。随后，他又从人的研究转向了猴子研究，之后再次转向了小鼠的研究。

在《通用人工智能》一书中，他如此总结自己的学术思想：如果只是沿着熟悉的道路前进，虽然容易且安全，却可能让你陷入认知的局部最优陷阱——你以为自己已经理解了整个世界，实则只是固守在一个狭窄的角落。

这就要不断迭代自己：在开始训练前，都必须明确一个目标函数。这个函数定义了模型希望达到的理想状态，而训练的全部意义就在于不断优化参数，让模型越来越接近这个目标。

人生也需要定义目标函数。他说，不用执着于精确地规划未来的每一步，那样反而可能陷入过度分析而迟迟无法行动。你需要做的是“强行起飞，粗糙开始，空中加油”，找一个大致正确的方向，然后不断调整，不停向前。

这正是刘嘉实现信仰之跃的路径。
(本文部分内容参考了《人文清华播客：对话清华学者》的刘嘉专访)

一位青年导师的成长与蜕变

■本报记者 杜珊珊

刘文丰与“水”结下了深厚缘分。

2010年，刘文丰进入北京师范大学，师从徐宗学攻读水文学及水资源的硕士学位。“徐老师曾说‘踏遍五洲皆为水’。他的足迹遍布全世界50多个国家，始终关注水文水资源科学研究与人类福祉。”刘文丰回忆说。

在徐宗学的引荐下，刘文丰结识了另一位对其科研生涯产生深远影响的前辈——瑞士联邦供水、废水处理与水环境保护所教授杨红。如果说徐宗学是刘文丰步入“水世界”的“引路人”，那么杨红则是助他夯实科研学术根基的“策路人”。

2013年，硕士毕业的刘文丰远赴瑞士，加入杨红研究团队攻读博士学位。“杨老师对每位学生都倾注了极大的耐心，时常强调严谨的重要性。”他至今记得，在修改其博士课题第一篇学术论文时，哪怕是一个看似不起眼的措辞，杨红都会反复推敲。这种“抠细节”的要求，进一步改变了刘文丰的科研写作习惯。

在杨红的指导下，刘文丰逐渐建立起系统性的科研思维和全局视角。“杨老师让我意识到，做研究不能只盯着局部问题，要始终保持对整体的把握，理解每一个变量在系统中的位置和意义。”

2020年，刘文丰结束法国气候与环境变化实验室博士后研究，回国加入中国农业大学，开启独立科研新征程。

在这里，他遇到了科研道路上的又一位重要“拓路人”——中国工程院院士、中国农业大学教授康绍忠。康绍忠长期扎根甘肃石羊河流域，亲赴一线带学生、做调研，专注于我国水资源与农

业水问题研究。“这让我懂得，科研做到这种境界即是艺术，更让我明确了科研初心：要把论文写在祖国最需要的地方。”刘文丰说。

现在，刘文丰的研究重心聚焦于气候变化背景下我国西北旱区水资源可持续利用这一关键课题。“在石羊河研究团队中，我们依托两大国家级科研平台，实验条件日趋完善。这份支撑更让我坚定信心，要将青春热忱挥洒在西北广袤的土地上。”刘文丰说。

科研的底色

如今，刘文丰已在中国农业大学任教近5年，组建了属于自己的课题组。

在刘文丰看来，科研培养的核心是点燃思维火花、唤醒问题意识。他不倾向于频繁的全体组会模式，更推崇灵活的小组讨论和一对一深度指导。“让不同方向的学生分小组推进，根据研究进展随时交流，而不是为开会而开会。”他说。

正因如此，课题组内逐渐形成了一种“轻结构、高效率”的协作方式。学生在各自研究方向深耕细作，遇关键节点便与他深入碰撞思路，形成开放研究的良性循环。

与此同时，他也格外注重培养学生的科研自主性。“如果有新想法，我会引导他们自主制定时间规划与目标节点。”在他看来，主动规划进度、明确攻坚方向，是成为独立科研工作者的必经之路。

刘文丰还有个特别的理念。他认为，博士生不



刘文丰

受访者供图

应是被动接受知识的“学习者”，要成为特定科研问题的“深耕者”。“进入博士中后期，学生在研究问题及技术细节上应比我更精进、专业。”他坦言，“这时我不再是单向输出的指导者，而要让他们具备‘反哺’能力。这样才能得到真正的成长。”

彼此信任、相互启发的相处之道，构成了刘文丰课题组更灵活自由、更尊重个体节奏的科研韵律。

“学生们常常带着新发现兴冲冲来讨论，有时一聊便是一下午，连饭点都忘了，但大家都乐在其中。”他笑言，“这种为了一个问题共同钻研思路、碰撞火花时刻，正是科研最动人的底色——它本就是让人心生动人的事业。”