

# 计算人文学：智能与意义的嵌入建构

陈跃红

**摘要：**人工智能的应用是一把“双刃剑”，无论在 AI+，还是 +AI 的领域，如何对其实施人文“审计”，使之实现有益于人类的利用是关键问题。目前紧要的不是对刚起步的人工智能应用做出“终极伦理”审判，事实上，我们更需要关注的是如何处理好人工智能发展过程中的一系列“过程伦理”问题，如何在它的发展过程中不断地建立法律和伦理的规训，推动人文价值和意义的关怀引导，从而促进人工智能技术这一新兴行业的健康发展。

**关键词：**AI 计算人文学 过程伦理

2018年5月19—20日，围绕人工智能的若干命题，一场题目为“人工智能时代的技术与人文：跨学科对话”的多学科交叉研讨会，在中国深圳最具创新活力的新兴大学南方科技大学召开。这次会议有一个显著特点，它既不是通常那种人文学者圈子内的形而上自说自话，也不是人工智能专家之间的纯粹技术交流，而是各自面向对方的一场认真严肃的跨学科对话。

会议邀请了18个报告人，其中9位是人工智能领域的专家，包括3位计算机领域著名的IEEE Fellow，另外9位是人文社科相关领域的知名学者，构成了一场特殊的9+9对话。

与会嘉宾在报告和讨论中普遍认为，人工智能研究和应用的各个结构性层面，都存在众多技术与人文密切相关的问题，这尤其需要技术专家与人文学者一起来探讨和解决。与此同时，与会专家也认为，人工智能是一把“双刃剑”，如何合理利用才是关键问题。目前紧要

的不是对刚起步的人工智能应用做出终极伦理审判，而更需要关注的是如何处理好人工智能发展过程中的一系列“过程伦理”问题，关注如何在它的发展过程中不断地建立法律和伦理的规训，推动人文价值和意义的关怀引导，从而促进人工智能技术健康发展。

鉴于此，建构发展一门专门研究人工智能相关人文问题的“计算人文学”学科，就显得十分迫切和重要。

围绕计算人文学的建构，需要简单回顾一下 AI 的历史。

事实上，有那么一些重要的人类历史的重要节点，并非都是在轰轰烈烈的仪式下宣告诞生，而往往是在平静的，不引人注目和不声不响中开始它的历史进程。可以肯定，1956年，当摩尔、麦卡锡等人在达特茅斯给 AI 命名的时候，他们并没有预料到，在 21 世纪的第二个十年，AI 会在地球人中间引起如此的轩然大波！同时另一方面，我们当下也越来越明显地感受到，现代历史过程中无数种技术进步和革新，譬如蒸汽机、发电机、电话、火车、汽车、飞机，等等，它们当中似乎没有一种技术像眼下的人工智能这样，才刚刚走向应用就引起如此多的巨大伦理争议，并且是前所未有地从价值、意义、伦理、法律、隐私、文化和美学等各种层面，与人文学科嵌合交集到了一起。人工智能作为一种历史性的技术进步，它在某种程度上超越了传统的，被动的工具理性，一跃成为具有主动性和自主性的新兴技术，它的“仿人类性”特征，使得它从一开始就不可避免地面临了如何跨越人文价值种种界面的挑战。

这是自工业革命以来雄霸世界的科学技术界所未曾预料过的局面。

同样我们也注意到，这一技术的应用性展开，几乎一夜之间就成了世界和各个国家具有战略意义的大事件。联合国、欧盟、美国、英国、日本、韩国等都发表了推动和规范人工智能快速发展的战略性国家报告。中国不仅发布了国家战略报告，还将之写进了两会政府工作报告和新的国家科技进步规划，足见其受到关注和重视的程度。需要特别指出的是，这些所有的报告，均有一个不同于以往任何重大科学技术

发展报告的特点，以前那些个报告基本上都集中关注科学、技术、人才、资金和政策层面，不太会，也不可能去讨论有关的人文问题。但是，眼下所有的这些人工智能发展报告，都无一例外的花费了相当的篇幅来讨论人工智能的人文社科命题，譬如法律、隐私、伦理、意义等。有的报告干脆就认为，人工智能领域天然的就是“游走于科技与人文之间的学科”，它既需要数学、统计学、计算机科学、神经学、脑科学……也需要哲学、心理学、法学、语言学甚至文学和美学。跨学科的研究方法和普遍的人文视角，眼下已经成为人工智能的重要方法学和研究进入角度首选。我国不久前发布的《新一代人工智能法发展规划》，一口气 15 次提及了“人工智能伦理”这一关键词。你不可想象其他任何科学技术发展规划会这样来强调科技与人文的关联，由此可见，将人工智能与人文学科加以整合，就其嵌合关联的现象事实和可能的结果展开系统研究，是学科交叉整合的重要对象目标，也是一种现实的学科研究需求，我们就不妨称之为计算人文学。

那么，计算人文学可以研究些什么呢？梳理一下人工智能研究和应用的各个层面，就会发现，无论是在基础设施层，算法层，还是技术及其应用层，都有众多的问题亟待解决。

譬如在基础设施的层面，即所谓硬件，算力和大数据领域，围绕数据的开放性、安全性、保密性、价值评估和资本意义等，就不可能是纯粹的技术问题，甚至在很大程度上应该归属于关于数据的价值认识论命题，数据的资本价值和人类共享权利张力的关系处理，数据使用和保护的建构，数据应用的正当性和道德正义边界的严格划定等等。Facebook（脸书）的数据泄露事件，扎克伯克的国会作证和道歉，“剑桥数据”公司的破产，不过是给世界敲响了一记警钟，9000 万被泄露用户的数据，被一类 APP 应用优选出 17 个州的中间选民的 32 种性格去推送广告，由此影响了大选的行情，而事情迄今并没有自此完结。特朗普的当选，英国不可逆转的脱欧，国际政治格局的骤然洗牌，在这些后面，曾经为“剑桥数据”设计选举推广算法的人工智能大师

作何感想我们并不知道，但是却令人思之恐极。

在基本算法层面，人文的思维逻辑，研究范式和经验知识，将有可能成为人工智能演进发展的精神源泉和内在动力。一些算法的特征和灵感来源，一些优化模式的改进和深化，都不仅仅是一种数理科学思维和技术自然逻辑的演进，而是与人类的形象思维，感性逻辑甚至生物生存规律现象的经验总结密切相关。譬如联想算法与反向传播，符号算法与逆向逻辑，进化算法与自然选择，类推算法与相似性测定等。当然，也许还应该算上在座姚新教授的演化算法的蚂蚁归巢比喻，史玉回教授的群体智能的群鸟寻食创意，这些算法和现象之间始终存在着内在的关联。

而在人工智能的具体技术方向层面，无论是图像识别与理解、语音的识别与合成、自然语言处理与机器翻译、情感分析、规划决策系统和各种大数据的分析，还是由此生发出的行业解决方案，如各种机器人功能岗位（例如仓库、港口、搬运、家居、超市、物业、书店等，金融、医疗、交通、旅游、区域安防、无人机和无人驾驶汽车……），机器在带来无与伦比的精准、快捷和便利的同时，人的生命过程“痕迹”将被无所不在的加以搜集、存储和精准操控的推送处理，而隐私、自由和随心所欲的个性化生活、多元化的审美文化建构，都有可能逐渐消失。如果说，在+AI应用的范式下，人工智能技术的使用正当性和合法性，尚有先在的功能设计和规范前提提供保障，那么，在AI+的范式下，其合法性就需要从一开始就必须加以严格的价值考量，法律规训和伦理评估，并且设计出既保障快速发展又能实施有效管控的系统规则，从而争取避免类似核武器出现的事后担忧和惊恐，这也正是我们以为有必要推动计算人文学这种跨学科研究的重要动因。

当然，说到底，AI也只是一种技术，一类程序、算法和优化手段，尤其我们目前基本上也还只是处在所谓弱人工智能时代，程序、算法和结果之间尽管存在运行的黑箱和晦暗不清的面目，但它本身目前至少还没有能力计算欲望、感情和意义追求，它的一些看似价值选择的

惊人之举，多数时候只是算法或者设计漏洞所致。因此，到目前为止的人工智能程序和算法，肯定不是，也不太可能是点石成金的神器或者是随时放出恶魔的潘多拉盒子，那些关于人工智能的终极伦理判断（譬如关于多大比例的人群将无所事事，多少人将毁灭或者永生的预言），在历史尚未终结的长时段内，谁也没法证明或者证伪。相比之下，我们也许更应该关注人工智能的“过程伦理”问题，尤其应该在 AI 的发展过程中不断地实施人文意义检测和伦理干预，目的在于推动它的健康发展。

如此，在人工智能应用爆发的当下，我们确实感觉到了文理工跨学科共创嵌合的现实脉动和互动对话可能，由此来理解耶鲁大学校长苏必德（Peter Salovey）所强调的“我们已经到了最需要人文学科的时候”，也许他并非信口开河。相比较机器会把人变成奴隶的预言，近代以来显示的事实却是：真正将人类变成奴隶的生物，从来都是人本身。

不过，迄今为止，我们对人类自我完善的能力依旧未曾丧失信心，我们对人类滑入命运深渊的预言始终存在质疑，我们对人类的前景尚存信心！作为 21 世纪拥有充分智慧和经历太多历史教训的人类，面对 AI 作为又一波技术飞跃的能量爆发，我们依旧能清醒的意识到，相对于预期中的强人工智能和超强人工智能时代，我们毕竟还处在弱人工智能的门槛之内，AI 的主要利器似乎还局限在算法和数据的逻辑互动关系层面，在逻辑智能的路上我们似乎已经走出了一大步，但是在情感智能和意义智能的路上，我们至今举步维艰，前景晦暗。当下人工智能在每一个领域的应用，都立即会爆发各种人文的、伦理的、法律的以及价值意义的质询，这应该看作是好事，是人类面对技术开始日渐成熟的一种标志。

显然，一切都还来得及，不过需要强调的是，身处人工智能这样一个科技与人文实质上已经相互嵌合的时代，我们不能只是单方向地，一味要求理工男们要抓紧脑补人文通识的知识，要培养人文精神和技术创新的价值灵魂云云。文理分科以及过度专业分工的恶果，并非只

是理工领域受害，人文学界也受害不浅，瞧瞧当下想走却又走不进人工智能技术逻辑场域，论述总是隔靴搔痒，学科圈子自嗨的大把期刊文章，的确不能不说，人文学者真的到了需要走出自己的知识“围城”和自我循环论证孤岛的时候了。在这样一个真正学科跨越和相互嵌入的时代，我们不能只是要求别人必须要人文补课，人文学界自己更加需要给自己来一场理工和技术的通识教育补课。即使仅仅是就发展计算人文学知识建构的迫切需要，我们也要奋勇走出书斋。走进实验室，走进 AI 技术世界的各个层面，然后才会获取到研究所需的知识 and 数据，结合自己的知识疆域和论证特长，取得真正有“技术含量”的发言权，从而真正从过程伦理的意义上参与这场对话，并且在双向互动中努力去建构起计算人文学的话语平台和对话机制。

商  
海  
印  
書  
館