数智时代的人文社会科学发展路径一探

邱泽奇（北京大学博雅讲席教授）

DOI:10.14167/j.zjss.2026.05.015

数智技术对人文社会科学（以下简称“文科”）的影响已然发生。素材在变，方法在变，研究对象在变，文科赖以成立的认知前提也在变。现在再讨论“是否接受”影响，已经没有意义，因为现实已经有了回答。进一步的问题是，我们接受影响的路径有几条？哪一条才是现实路径？沿着这条路径走下去，文科又会有何新发展？

一、文科范式变革的现实

文科曾经历过两轮范式革命。第一轮是学术革命，以专业化建构了现代人文学科。文学、史学、哲学等学科得以从笼统的人文学问中分立出来，各自发展出研究对象、方法与论证形式。第二轮是计量革命，以计量方法建构现代社会科学。社会学、经济学、心理学等学科从此把可观测、可计量、可检验的概念与命题作为研究的核心环节。

数智技术正推动文科进入第三轮范式革命。这一轮可划分为素材像素化、像素数据化、数据理论化、理论智能化等四个阶段。（邱泽奇，2023a）[①]当前的发展主要处于前两个阶段，第三、四阶段也已启动。四个阶段交错并进、相互推动，构成当下文科范式革命的真实格局。素材像素化把文献、图像、声音、行为都纳入可处理的数字形式。像素数据化把这些原始形式整理为可被算法读取的数据结构。数据理论化通过算法在数据之上生成对社会现象的概括与命题。理论智能化进一步把已有理论植入机器智能，让理论成为可被机器运用与迭代的资源。前两个阶段是基础设施建设，后两个阶段是知识生产能力的更新与文科存量知识的再激活。新的范式不会一次性替换旧的范式，旧研究路径的产出仍是新范式有效的输入。学科边界、问题清单与论证标准，都在交错并进中逐步改写。

数智时代，多技术的融合让素材汇集与方法体系化获得了前所未有的有利环境。文献被大规模数字化，思想观点也具备了可计算性。文科研究第一次得到大规模、跨语种、跨时段的素材支撑。大语言模型对文献的语义索引、概念聚类与跨文本比照等，都在以个案研究和少数文本或素材无法达到的尺度上扩展。“定性—定量—计算”的混合方法已能在同一项目中并行运行，不同流派的拼接让位于一体化的研究设计。（张茂元、邱泽奇，2024；王国成，2025）

国际学术界也走在同一方向上。现代文科从实证主义起步，经过解释主义、混合方法、数字方法，最终走向人工智能中介的复杂研究。加布里埃尔·普齐亚诺（Gabriella Punziano）把这一过程描述为“适应性认识论”的形成。社会科学的研究范式正在向接纳新一代人工智能的方向调

①近三年笔者围绕数智时代人文社会科学的发展形成系列研究(详见参考文献)。本文涉及具体概念时单独标注,其他论点不再逐一引证。

整，研究伦理、方法选择与解释逻辑都进入新的组合。（Punziano，2025）克鲁斯-阿吉拉尔（Manuel Alejandro Cruz-Aguilar）在更宏观的层面认为，人工智能正在重组科学知识的认识论根基。“什么是知识”“谁有资格生产知识”“知识如何被验证”等根本问题，都在得到新的解答。（Cruz-Aguilar，2025）数据规模与算法效能的迭代，已经把过去依赖少数专家直觉的判断变成可被批量生成、批量检验的命题。文科研究的素材边界、方法边界与认识论边界，正在同步变动。

数智时代，还在让机器智能不再是被动工具。大语言模型与算法已具备语言理解、信息整合、判断生成等能力。它们正多路径地进入人类社会的核心领域，参与知识生产与社会治理。这意味着文科从超脱于对象的研究主体变成了同时是“被研究者”和“参与研究者”。一方面，机器智能作为新的社会力量成为文科要解释的对象。它如何介入劳动、教育、医疗、政务等关键场景，如何重塑社会团结的形态，如何影响公共理性的运作，都是文科必须回答的问题。另一方面，机器智能已经在支撑文科发展。在文献检索、概念聚类、命题生成、文本生成的各个环节，机器都能介入。研究的执行者与研究的客体在同一项目内交叠，这是之前从未有过的现象，也是人机互生的关键证据。

与此同时，人类认知的独立性与独特性正在失效。机器智能与认知科学的相互推动让认知场景始终处在陌生化状态。机器既支撑认知增强，也提供认知干预的机会。（邱泽奇，2023b）文科的一项前提曾经为“人是认知的唯一主体”。随着机器具备语言理解、判断生成与价值表达的能力，把人类预设为认知世界唯一主体的假设也难以为继。认知干预的形态包括算法推荐对注意力的引导、生成内容对记忆与判断的影响、个性化推送对偏好的反向构造。这些干预不一定带有恶意，却在事实上参与了人类认知的形成与改变。人工智能写作助手对学术文本的间接介入、对话式系统对用户语言习惯的潜移默化，都属于同类机制。它们不需要任何强制就能改变人们对世界的看法。文科若仍把人类认知视为不被外部介入的纯粹过程，可能无法处理如今大多数与人有关的问题。

面对如此变局，文科若不主动应对这一现实，有可能逐步丧失刻画与解释新现实的能力，从而动摇其存在的依据。复旦大学的《人文社会科学智能发展蓝皮书》提出，人工智能赋能下的文科进入了数据与机理双驱动的“第五范式”。（陈志敏、吴力波主编，2025）这与我们前面的判断是同一事件的两种表述，都把当下的文科放在一个不能回避的时代环境中。要么承认现实，并在现实中顺应研究范式的变革。要么留守旧范式，让自己的解释力一点点被现实抽空。当然，旧范式的研究路径不会立刻失效，可它的解释力会逐步收窄乃至失去。随着机器智能生成知识的能力不断增强，人类知识的合法性也将由人类对机器的引导能力决定。由此，学科科学革命以来的边界不再自动被赋予合法性，而必须由文科主动去重建。

数智技术对文科发展的影响已然发生，文科要回答的问题已转向接受影响的方式。这种转向有问题清单的更新，更多的则是文科自我定位的再造。文科要回答的问题扩展到了“人是什么、社会是什么”之外，“人与机器之间的边界在哪里、谁来设定这条边界、按什么原则设定”，已成为同等重要的新问题。

二、文科发展路径的选择

接受数智技术的影响有两条单极路径。一条跟着机器的效率走，一条跟着人类的福祉走。人

工智能哲学的讨论也提出了两种选择，温和式增强或激进式替代。（潘斌，2024）笔者以为，两条路径都不现实。现实的路径是把机器效率置于人类福祉的框架内，实现人机互生。

第一，跟着效率走，结果是人类的机器化。机器是效率工具的产物，沿着效率的方向继续演进是它的内在逻辑。文科若不为机器设定价值方向，工具理性就会垄断人类行动空间。新一代人工智能与生物技术的合流甚至会催生人类物种的分化。一旦“物种资本”成为社会达尔文主义的核心，生物人类的社会性承认就会被动摇，（邱泽奇，2025a）包括自信、自尊和自我实现。纯技术方案与成文立法都难以应对人工智能伦理风险的深层性与复杂性。（樊春良、吴逸菲，2025）企业利益与公众价值的冲突、技术黑箱、人类伦理共识的缺失，构成提升效率的持续动力。把方向交给效率，等于把价值的最终回答外包给市场与算力。在这条路径的尽头，文科已无位置可言。

第二，跟着人类福祉走，却忽视机器效率，结果是机器的人类化。这一路径无视人性对效率的追求，也忽视机器对人类的积极价值。它把文科退回前数智时代的舒适区或之前，甚至退回到宗教统治时代，本质上是对现实的放弃。文科以人为本，不等于以拒绝机器为标志。把人类的自我守护建立在对机器能力的不承认之上，结果是文科自身从现实中抽离，失去回应现实的地位。如今，已经没有哪一项重要的文科议题（劳动、教育、医疗、社交、治理）可以绕开机器智能。拒绝机器，最终伤害的不是机器，是文科本身。承认机器具有创造力，是文科主动应对数智时代的前提。（徐飞等，2026）

人性的两面，是效率追求与主体性坚守并存。两条单极路径都失败的原因是各自偏废了人性的某一面。人类既要更高效率的物质条件，也要可被尊重的精神空间。这两面都是人性，并非可以二选一的偏好。文科若只承认其中一面，就会替人类做出错误的选择。把效率拱手让给机器，等于让出物质生活的主导权。把主体性拱手让给传统，等于让出精神生活的丰富性。两种让渡都不可接受。文科要做的，是把效率与主体性放在同一个框架里。

人机互生是现实的第三种选择，它把机器效率置于人类福祉的框架内运行。人机互生有三种机制。一是人机主体的相互建构。人赋予机器的能力边界与价值底线，机器反过来影响人类的认知方式与行为模式，两者在持续互动中互为前提。二是行动的伦理协同。机器在具体场景的每一次行动，都需与人类的伦理预期保持协同，单方面的优化不构成可接受的行动。三是治理的实践递归。治理规则不可能一次性完成，需要在实践中不断被反馈、修正、迭代。（邱泽奇，2025c）三种机制合起来，把人机互生从一个口号变成可被研究、可被设计、可被治理的对象。目标是，在承认机器自主性的前提下，维系人类的主体性与人性的尊严。从社会团结的角度看，连接机制与团结形态也都在转型。社会连接机制从抽象连接走向人机连接。团结形态也由有机团结、生态团结走向人类团结。新的团结形态接纳人机共在，不再以“人之为人”的纯粹性作为团结的边界。人类社会的如此转型是文科之前未曾遇到的，也是如今文科发展的入口。人机互生是把人和机器放在一起，落到学科建设、研究设计与人才培养中的实操命题。

劳伦·克莱因（Lauren Klein）等从人文学科角度提出了三点提醒。模型造词、人造意义以及规模并非答案。解决之道有三个方向——承认编码视角、收敛于领域专属与小型化、抵抗算力垄断。（Klein et al.，2025）这从另一面支撑了人机互生的命题。机器需要意义，意义只能从人类来。在意义供给者位置上，文科具有不可替代性。让规模决定一切，等于让算力决定意义。让算力决定意义，等于让市场决定价值。这条链条的尽头，是人类把意义的最终裁定权外包给少数算力垄断者。这是世界性的人文学科警告，也是中国文科必须回应的议题。中国文科的本土资源与

制度安排恰好为应对这一警告提供了不同于西方路径的回应空间。

三、文科发展的崭新内涵

数智时代的文科有了新的内涵，一方面它为人类提供精神食粮，一方面它为机器价值对齐提供方向。两面一起，共同锚定人类福祉。数智时代文科的这条发展路径是同一个价值锚定的两面表达。为人类提供精神食粮，是对内的价值供给；为机器价值对齐提供方向，是对外的价值供给。两面合一，才真正构成了文科完整的发展路径。

数智技术已进入人类意义建构的过程。“灵肉分离—灵肉相融”的转换正在发生。（邱泽奇，2023c）随着人类的劳动、情感、记忆、社交都被数据化、算法化、平台化，意义的来源不再自动生成。意义感的锻造已经成为文科不可推卸的社会责任。在数智条件下，文科要再次回答“人为什么活着、什么是值得过的生活、人与人之间连接的价值在哪里”等问题。而要回答这些问题，文科自身的语言、概念与论证形式都要随之更新。

不仅如此，与机器社交意味着机器智能介入了人生意义的建构。为此，也亟须积极倡导和推动人工智能为善。（邱泽奇，2024a）机器只能在巨大的设计空间里选择目标函数，而函数的取向，归根到底得由人类社会的价值排序决定，文科是这一价值排序的提供者。机器价值对齐的源头，是通过人文社会科学研究，实现人为机器立法。（闫坤如，2024）仅靠技术路线，价值对齐易陷入先验理性的困局。价值对齐须走向人机协同的理性，让多元利益相关者在多样情境中动态参与。（宋保林，2025）问题是，既有的价值对齐方案多基于人机二元对立的先验架构，遮蔽了人工智能技术内部携带的价值倾向，甚至导致价值殖民。（吴静，2025）文科理当介入对齐的完整过程，提供持续的价值校准。脱离价值的算法不可能自我对齐，脱离文科的对齐工程会迷失方向。

还有，没有了人类，谈不上文科，也谈不上文科发展。人类存在的前置性回应了对工具理性垄断的警告，也回应了“以人为本、智能向善”的治理要求。（程乐、龚煊，2026）人类福祉是文科一切其他议题的前提，也是文科与机器智能对话的参照系。在数智时代，任何不以人类福祉为锚点的研究，都偏离了文科应当承担的社会责任。

运用数智技术的赋权与赋能，为人类福祉创造精神食粮，为机器效率的提升提供价值方向，同时回答了文科的工具更新与知识范式转变两个问题。文科智能是工具更新与知识生产方式更新。（邱泽奇，2025d）专家知识数据化是其前提与抓手。（邱泽奇、李昊林，2025e）从专家头脑里的隐性知识到可被机器使用的显性数据，是文科参与机器价值对齐的工程接口。具体的工程接口包括三类：一是把学者头脑中的判断标准转写为可被机器学习的标注数据；二是把学科共识的概念体系转写为可被模型使用的本体结构；三是把研究方法的隐性规则转写为可被流程复用的工作模板。文科不进入这一接口，文科与机器之间的对接就只是技术的“借用”。文科主动建设这一接口，文科与机器的关系就成了双向的协同。

当然，文科的新内涵并非凭空而立。在精神方向上，有清晰的顶层指引。2016年5月17日，习近平总书记在哲学社会科学工作座谈会上就提出过明确要求，哲学社会科学必须坚持以人民为中心的研究导向，“坚持为人民服务、为社会主义服务方向”[①]。这是为人类提供精神食粮的最高

①参见《习近平：在哲学社会科学工作座谈会上的讲话（全文）》，全国哲学社会科学工作办公室网站，https://www.nopss.gov.cn/n1/2016/0519/c219468-28361739-4.html。

政策表述。2022年4月25日，习近平总书记指出“加快构建中国特色哲学社会科学，归根结底是建构中国自主的知识体系”。[①]融通古今中外资源、坚持文化主体性，为文科在数智时代的本土锚点提供了方向。在2023年6月2日的文化传承发展座谈会上，习近平总书记进一步强调中华文明具有突出的连续性、创新性、统一性、包容性、和平性，提出建设中华民族现代文明。（习近平，2023）这一表述为机器价值对齐补足了文化根脉。2025年4月25日，习近平总书记在中共中央政治局第二十次集体学习时强调，“推动我国人工智能朝着有益、安全、公平方向健康有序发展”[②]。这把“为机器价值对齐提供方向”从学界呼吁上升为国家方向。四次表述层次递进，从研究导向、知识体系到文化根脉再到技术方向，勾勒出数智时代文科新内涵的精神坐标。

精神方向需要制度落地。国务院2017年印发的《新一代人工智能发展规划》明确把人工智能与社会融合、伦理治理、人才培养纳入国家战略部署。[③]这为文科介入人工智能议题打开了制度空间。教育部新文科建设工作组2020年发布《新文科建设宣言》，提出守正创新、价值引领、分类推进等基本原则。[④]其中价值引领作为新文科建设的方向，与为机器价值对齐提供方向同向。中宣部、教育部2022年联合印发的《面向2035高校哲学社会科学高质量发展行动计划》进一步把自主知识体系建设、新文科推进与人才平台建设落到具体实践安排中。[⑤]中共中央、国务院印发的《教育强国建设规划纲要（2024—2035年）》提出更具体的部署，要求构建以各学科标识性概念、原创性理论为主干的自主知识体系，部署深化新工科、新医科、新农科、新文科建设，强化科技教育和人文教育协同，把打造人工智能教育大模型、强化数据安全、强化人工智能算法和伦理安全作为支撑条件。[⑥]这些制度安排组合起来，为文科智能搭起了完整的制度脚手架。教育部等五部门2026年4月印发的《“人工智能 + 教育”行动计划》则把方向落到具体行动上。[⑦]计划明确坚持育人为本、素养为先、应用导向、智能向善，把哲学社会科学纳入人工智能赋能科研的主体范围。在具体动作上，部署研发思政大模型，强化价值对齐、逻辑推理、安全伦理等能力建设。在原则上，要求科技教育与人文教育相结合、技术应用与人文关怀相统一。从规划到行动，从战略到落地，制度保障与精神方向已经合流。五份文件层层推进，从首个国家级人工智能发展规划到2026年的“人工智能 + 教育”行动计划，文科介入数智议题的制度环境已经构建。剩下的工作是文科如何在有利的环境里推动自身的快速发展。

把精神方向与制度保障合起来看，中国哲学社会科学的自主发展（邱泽奇，2025f）与数智时代文科的新内涵，是同一事物的两面。一面是对主流大模型保持西方中心立场的回应，（陈志敏、吴力波主编，2025）另一面是对“数智时代文科凭什么立得住”的中国回答。把两者结合起来，文科发展就有了中国语境的具体落点，那就是，以自主知识体系为骨干，以人民福祉为价值

①参见《习近平在中国人民大学考察时强调 坚持党的领导传承红色基因扎根中国大地 走出一条建设中国特色世界一流大学新路》，人民网，http://jhsjk.people.cn/article/32408562。

②参见《习近平在中共中央政治局第二十次集体学习时强调 坚持自立自强 突出应用导向 推动人工智能健康有序发展》，人民网，http://jhsjk.people.cn/article/40468944。

③参见《国务院关于印发新一代人工智能发展规划的通知》（国发〔2017〕35号），中国政府网，https://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm。

④参见教育部新文科建设工作组：《新文科建设宣言》，中国社会科学网，https://www.cssn.cn/ztzl/yingxiangshixue/zhengcewenjian/202302/t20230216_5588800.shtml。

⑤参见《中共中央宣传部、教育部联合印发〈面向2035高校哲学社会科学高质量发展行动计划〉》，中国政府网，https://www.gov.cn/xinwen/2022-05/27/content_5692590.htm。

⑥参见《中共中央 国务院印发〈教育强国建设规划纲要（2024—2035年）〉》，中国政府网，https://www.gov.cn/gongbao/2025/issue_11846/202502/content_7002799.html。

⑦参见《教育部等五部门关于印发〈“人工智能+教育”行动计划〉的通知》（教科信〔2026〕1号），中国政府网，https://www.gov.cn/zhengce/zhengceku/202604/content_7065138.htm。

锚定，以人机互生为运行机制。以自主知识体系为骨干，意味着文科以中国问题、中国经验、中国方法生成适用于人类的概念与命题，避免对西方理论概念体系的单向依附。以人民福祉为价值锚定，意味着文科的研究目标始终把人的尊严与发展放在第一位，避免被市场效率与技术效用垄断和定义。以人机互生为运行机制，意味着文科与机器智能形成合作关系。文科作为价值的供给者，机器作为效率的实现者。两者共同选择了中国文科发展在数智时代的发展路径，既不同于让文科退回旧范式的保守取向，也不同于把文科交给大模型的被动取向。

在数智时代，文科的未来，将通过人机互生的知识创新与积累，成为人类日常需求的知识源。（邱泽奇，2024b）这是数智时代文科发展路径的最终去向。沿着这条路径走下去，文科将以两个新身份立于机器智能时代：一是机器智能价值方向的供给者，二是人类福祉的守护者。两个角色的融合，构成数智时代文科存在与发展的根本理由。

参考文献：

1. 陈志敏、吴力波主编：《未来已来——人文社会科学智能发展蓝皮书》，复旦大学出版社2025年版。

2. 程乐、龚煊：《人工智能治理的逻辑思辨与人本主义路径》，《浙江大学学报（人文社会科学版）》2026年第2期。

3. 樊浩、邱泽奇等：《意义感：数字时代民众的幸福感知与价值锻造》，《探索与争鸣》2023年第6期。

4. 潘斌：《增强还是替代：人工智能与人类智能关系的哲学反思》，《理论月刊》2024年第7期。

5. 邱泽奇：《数字化与文科范式革命》，《大学与学科》2023年第4卷第2期a。

6. 邱泽奇：《认知域：从习以为常到人机互生》，《人民论坛·学术前沿》2023年第11期b。

7. 邱泽奇：《从灵肉俱在到灵肉分离再到灵肉相融——数字时代日常意义的个体汇聚》，《探索与争鸣》2023年第6期c。

8. 邱泽奇：《机器智能与社交关系重构》，《人民论坛·学术前沿》2024年第19期a。

9. 邱泽奇：《文科智能的未来图景》，《人民论坛·学术前沿》2024年第2期b。

10. 邱泽奇：《人类物种分化及其主体性的未来》，《人民论坛·学术前沿》2025年第22期a。

11. 邱泽奇：《人机社会的治理转向》，《社会》2025年第3期b。

12. 邱泽奇：《数智时代的人类团结——技术变革与理论重构》，《智能社会研究》2025年第1期c。

13. 邱泽奇：《人机互生：范式革命与知识生产重构》，《社会科学文摘》2025年第6期d。

14. 邱泽奇、李昊林等：《专家知识数据化与文科智能》，《中国社会科学》2025年第7期e。

15. 邱泽奇：《数智时代中国哲学社会科学自主发展的机遇与挑战》，《光明日报》2025年8月15日f。

16. 宋保林：《人工智能大模型价值对齐的伦理建构》，《伦理学研究》2025年第3期。

17. 王国成：《数智技术与人文社会科学的有机融合》，中国社会科学网，2025年6月17日。

18. 吴静：《价值对齐能否解决人工智能风险》，《中国社会科学评价》2025年第4期。

19. 吴逸菲、樊春良：《中国语境下人工智能伦理治理的路径塑造——基于价值—工具二维理性的融合框架》，《科学学研究》2025年第1期。

20. 习近平：《在文化传承发展座谈会上的讲话》，《求是》2023年第17期。

21. 徐飞、董小龙、何晓松：《人类如何面对人工智能的创造力——从怯于接受到人机共创》（网络首发），《科学学研究》2026年4月3日。

22. 闫坤如：《人工智能价值对齐的价值表征及伦理路径》，《伦理学研究》2024年第4期。

23. 张茂元、邱泽奇：《数字社会研究框架与思路》，《清华社会学评论》2024年第21辑。

24. Cruz-Aguilar, M.A., "The Epistemic Revolution of AI: Reconfiguring the Foundations of Scientific Knowledge", *AI & Society*, 2025, 41(3): 2041–2057.

25. Klein, L. et al., "Provocations from the Humanities for Generative AI Research", arXiv: 2502.19190(2025).

26. Punziano, G., "Adaptive Epistemology: Embracing Generative AI as a Paradigm Shift in Social Science", *Societies*, 2025, 15(7): 205.