

建立一个“人工智能符号学”

◎赵毅衡

摘要：人工智能的超速演变，不久就将替代人类智能主导当代文化。这个人类历史上空前未有的巨变，正在我们眼前一步步发生，已经成为当前世界最重大课题之一。对于人工智能的发展前景，我们可以用思想实验的方式，建立一个“人工智能符号学”，便于真切理解。从符号学机制看，人工智能进展速度快，正在迅速耗尽文化意义活动的基本动力，即认知负熵，人类文化活动正快速奔向“意义热寂”。因此人文学科应当警告世人：人工智能很可能将给人类文化带来重大危机。在拥抱人工智能时必须加以预防，才能让人类避开灾难。

关键词：人工智能；认知差；熵增；意义热寂；人工智能符号学

DOI:10.19290/j.cnki.51-1076/i.2024.06.005

引言 文本的根据与目的

符号学能否用来探索人工智能对人类文化的影响，尤其是预测其前景？本文坚持符号学可以处理这个问题，因为符号学是意义形式之研究，符号与意义是一体两面。而人工智能在人类文化中的角色，就是意义的产生，传送，解释，回应，以及进一步展开。整个文化是社会相关符号行为的总集合，正如麦克卢汉所说：“文化总是体现为各种各样的符号，……文化的创造在某种程度上说就是符号的创造……所谓文化，究其本质乃是借助符号来传达意义的创造、交往、理解和解释。”^①因此，

符号学用来检查考察人工智能的文化后果，非常适用。

但是这问题，不该只由我来回答，或许应当由人工智能来回答。人工智能是否了解自己在做什么？这正是本文要探索的问题。目前让人惊异其能力的各种人工智能系统，似乎无所不能，它们对各种问题的迅疾回复，只是该系统的数据库中产生连接的资料，至今尚不是该系统作为解释主体“经过深思熟虑”的回答。但是人工智能系统对自身问题的回复，可以作为本论文讨论的出发点。

正因如此，如果读者对本论文的论据有所质疑，不妨自己动手，与各种人工智能系统“切磋”一番，无论得出什么结论，都欢迎商榷。

第一个问题：人工智能的基本功能是否为意义活动？产生的是否为符号文本？能不能用符号学来考察？

很多人工智能系统的回答骑墙，说“有可能是意义活动，有可能应用符号学”。但 ChatGPT-4o 给了非常明确的回答：“人工智能的基本功能可以视为意义活动，产生的确实是符号文本。符号学，作为研究符号及其意义形式的学科，完全可以用来考察和分析人工智能的运作和输出。”既然回答如此明确，我们就从这个前提，开始讨论。

的确，人工智能的基本功能，就是对数据（即资料 Data）的处理和分析，数据库里的是文本、图像、声音等，无论数据是哪种形式，都符合符号的定义：“被认为携带意义的感知”。也就是说，人工智能的数据库，就是一个巨大的“符号库”，只是符号尚未被选中并发出，因此数据库是个“潜在符号”的大水库。人工智能系统的“深度学习”就是把这符号水库做大，可以从中找出对几乎任何问题的意义连接方式。人工智能系统的输出，无论是回答问题、生成文章、进行对话、产生图像、视频或乐曲，抑或是翻译成各种语言、输出各种指令，都是在传送出携带意义的“可感知”的符号文本。

需要说明的是：“人工智能符号学”（AI semiotics）是研究人工智能运作与其文化影响的方法。有不少人工智能系统本身（以及某些人类研究者）望文生义，把这问题混同于人工智能设计研究中的“符号主义”（symbolicism）。符号主义是人工智能研究史上一种方法论，源于数理逻辑，认为智能产生于抽象思维，如数学推导和概念化。它强调符号之间的逻辑关系和符号操作的规则。与之形成对比的是人工智能研究中的另外两种主要指导思想，是模拟大脑神经元之间的相互作用及信息传送原理“联结主义”（connectionism），以及模拟心理学与控制论，以可观测的具体的行为活动为基本工作方式的“行为主义”（actionism）。目前在人工智能设计理论中，这几种方式在进行符号意义活动，正如我们在观察高等动物或幼儿的符号认知活动时，也能发现这几种方式并存。

而本文所论的“人工智能符号学”，不是探讨

人工智能内部设计与结构方式，而是考察各种已经形成的人工智能系统，在人类文化中的表现，考察它们对人类文化可能产生的影响，简单一句话，就是“人工智能符号学”做的是人文学科式的“外部研究”，而“符号主义”等是人工智能的“内部”算法研究。把内外二者相混，没有任何好处，只能把问题说糊涂。

“人工智能符号学”作为一种学术探索方向，目的不是介绍人工智能发展过程，也并非意图普及人工智能的工作原理。人工智能对于绝大部分知识分子（包括人工智能从业者研究者）是个黑箱。“人工智能符号学”不必进入这个黑箱，其目的是在人文学范围内促动一种对人工智能的反思能力。人工智能符号学将从人类文化角度理解这场沧海巨变对人类的意义，看它正在把人类社会引向什么方向。

人工智能使用深度学习算法，模拟人的神经网络，来寻找数据中的优势连接。这一过程从“预训练”人工智能开始，将大量数据输入，使其能够找到这些趋势，沿着这一趋势线，推断生成新数据。因此，它们做的是大数据输入，按关键词查询连接，形成符号文本输出。不管是计算机、自动驾驶汽车、智能无人机、聊天机器人，虽然输出文本及其媒介大相径庭，做的工作却相似。

让人工智能系统“自白”，是建立一个人工智能研究的最好的起点。既然人工智能系统号称一个“智能体”，就应当明白自身的是如何存在于世的。人类历史过了多少世纪，才有意探问自己究竟是什么样的一个存在？对“我在”的理解，开始了人类文化现代化的进程。就此而言，我们的这次研究，给我们希望，更让我们警觉：人工智能或许尚未明白它们在做什么，人类却必须明白人工智能在做什么，现在就弄明白。

回答本文有关问题的人工智能系统，有《百度文心一言 4.0》《阿里通义千问 4.0》《灵犀人工智能 GC 系统 4.0》《天工人工智能 3.0》等。这些都是好学生，回答得头头是道。但也正如“做题家”的风格，回答大同小异，严重缺乏个性。既如此，我们的分析就暂时不去讨论它们的差别，也不讨论它们

重复中的差异，有兴趣的读者可以自行获取。我们可以检验它们“能否像人一样思考”，尤其是有没有可能获得人的“心智理论”ToM（Theory of Mind）方式。

讨论的问题，前五个分别询问人工智能的“符号处理能力”，以及与符号学研究的相关性；后五个问题，则是有关人工智能发展会对人类文化造成何种后果。因此，人工智能对人类文化的影响，是一个“AI符号学”探索的重点。

第二个问题：人工智能依靠数据库的资料，汲取的元素是信号还是符号？也就是说，这个过程是否包含着符号的理解和解释？

首先说清：什么是信号（signal）？它与符号有什么不同？信号是一种落在门槛上的符号类型，信号携带着清晰的意义，看起来完全符合“符号”的定义。但它不需要解释，而是要求预定的实际反应。动物的符号大部分是信号，对信号，动物一般不解释，而是依照物种的生理机制预先安排好的方式，采取固定行动。例如野羊逐草而居，鲑鱼群溯流而上。动物植物之间、身体内部器官之间都会有信号。并不需要（也不允许）接受者的思考，来决定反应方式。机械之间的反应，也是基于某种信号，例如电梯门关闭时，因光被遮断而重新开启。

那么人工智能在进行上一节所说的各种意义活动时，究竟有没有做解释？例如应对“提示”（prompts）进行回答，显然，人工智能系统给出的，不完全是信号的固定回应，而是做出一定解释的。用户的提示往往是语言符号文本，人工智能使用预先训练好的大语言模型，来解码这些输入，理解其意义结构。它们通过统计模型、语义分析、语法规则等，对有关的数据进行处理和组合。最终呈现给用户的，是各种符号组合文本，例如图像、视频、语言或文字的报告。在整个过程中，人工智能并不仅仅是处理信号，而是在解释符号。各种系统的运作确实包含了一系列的解释，人工智能使用的，不是简单的信号传递与反应，而是符号意义再构建。

然而，既然人工智能是一种“人工智能”，也就是“非人类智能”，我们就必须了解它与人类的

思考方式究竟有什么异同？可以看得出，各种“文生内容”（AIGC）系统的回答相当机械，缺乏个性，在符号的“三分式”，即再现体-对象-解释项的三联中，解释项明显缺席。“对象”是“文化中规定的连接”，是外延式的回答。可以发现大部分人工智能系统很善于给个不痛不痒，左右逢源的回答。它们经常像一个学习勤奋的大学生，文献综述做得相当不错，但是拿不出自己的眼光，无法在关键问题上清楚地表明立场。正是在独立见解方面，它们的表现远远赶不上人类。人的智慧不在于面面俱到，而在于明确地说出自己见解的判断，并用自己的思索加以论证。就它们对问题的解释而言，至今看起来，人工智能各系统依然像是一个“搜索引擎”的升级加强版。

第三个问题：人工智能的响应的过程是不是一个聚合操作？什么元素被选下？

人工智能选取答案的过程，的确是一种有确定方向的聚合轴选择操作，是基于相关性和概率的选择过程，目的是给出一个可行答案。人工智能通过模型的运算，决定哪些数据可以最终构成输出文本。这一步涉及对大量可能的候选项进行评估，以确定哪些符号最具有相关性。此种评估基于预训练（pre-trained）模型，只要数据库足够大，就能找出输入的提示所要求的答案。

人工智能会根据特定的标准，对这些候选项进行筛选和过滤，通过评分机制来选择最优的几项。经过聚合轴的筛选后，得到最优的符号组合，推出最终的输出文本。所以人工智能有选择，其标准是相关性和概率：相关性决定了与输入提示匹配，而概率评估则在可能的符号组合中找出“最佳”选项，也就是数据材料中的最佳答案。

有的专家认为，人工智能对数据库素材进行质量评估，选择来源可靠的材料。但也因此缺乏针对用户需要的个性。但是至今人工智能已经开始能做到考虑用户偏好：基于用户历史行为模式选择，选择最适合用户需求的材料。我们可以相信，人工智能不久就会适应“熟客”的需要，寻找客户需要的，迎合客户偏好的答案。此种个性或棱角，或许会引用户进入“信息茧房”，夸大其词，制造偏听

偏信。此时我们反而会怀念目前“直来直去”的诚实人工智能系统。

如果说人工智能的回应严重缺乏原创性，这是可能的。不过机器的原创，也是很危险的前景：如果人工智能有能力设计出骗人骗己的偏见选择，人工智能就将失去解释的客观性，这将让用户面临智能不可信任尴尬局面。

第四个问题：人工智能自身会不会犯错？如果数据错了人工智能能识别吗？错了如何纠正？

人工智能会犯错，但往往是数据中存在偏差或错误，尤其是在实际应用中，输入数据可能包含不相干的冗余符号，即噪音。要识别数据库中的错误，至今人工智能的能力有限。上一节已经说过，目前人工智能欠缺对数据质量与对错的评估能力。据说某些先进的人工智能系统，可以检测数据中的异常值或异常模式；通过多源数据交叉验证，识别可能的错误数据，进行一定程度的纠错处理，但离普遍的测错机制还很远。或许在今后在训练过程中，通过反馈和迭代优化，新的系统能逐步提高识别错误数据的能力。尽管如此，要人工智能在数据库的现有连接中识别对错，至今尚无可行道路。

目前的人工智能各种系统，不会“意识到”自己产生了错误回答，更不会发觉回答中不够完美的地方。只有在用户不满意，进一步追问时，它们可以在新的提示基础上提出一些补救。如此通过事后反馈补救错误，当然是很不够的。某些错误（例如自动驾驶）因为来不及补救，将是致命的。

人类追求真相时，用的符号学逻辑方法，是所谓“试推法”（abduction）。前提是明白：任何进行符号意义活动的人，不可能避免错误，而是明智地不固执己见，不断纠正错误使意义活动向前推进。如果人工智能不能克服这一关，也就是说不能自我发现错误、自我更正，那么就离具有评估能力的“主体性”差得很远。例如它只是一个一天能写出上万首诗、却无法自己选出几首给大家读的“电脑诗人”。

第五个问题：人工智能能否用迂回的办法对产出文本进行修辞，例如比喻、象征、反讽、悖论？

这些“有意说反话”加强修辞力量的意义方

式，是人类意义活动最简单的方式，却也是对“人工智能是否具有人性”的最有说服力的测试。人的意义活动，经常不直接按字面处理，而是根据交往惯例，根据符号使用的语境，根据对交谈者的性格的理解，来判断某些话究竟是什么意思。修辞，即拐着弯说话。有时是因为无法直接说，更多的是为了使表达生动委婉，目的是激活解释对字面下隐含意义的理解。

对各种人工智能系统进行这种测试，很简单易行。很快我们就可以发现两个倾向：一是对提示中的各种修辞手法，人工智能大致能够感到并指出提示语句中有语义矛盾，或意义模糊之处。但是人工智能自己不会主动采用曲折修辞手法，所以人工智能生成文本多半很刻板。一是人工智能可以看出“提示”里话中有话，问题中语义有矛盾。它们的回应多半是给出二层回答：你可能是这个意义，你也可能是另一种意思。例如问话：“你那么喜欢打游戏，就别回家了”，人工智能体系的回答大致上是：“你可能是生气了，但你也可能是认真的。”有的系统回答：“适当游戏有益于身心健康，但是回家机会也很宝贵。”甚至家长式的劝导：“善于游戏并不意味着您会因此失去方向感或忘记回家的路。我们应该将游戏和现实生活区分开来，并在享受游戏带来的乐趣的同时，保持对现实生活的关注和责任感。”此种刻板往往令人啼笑皆非。

例如，提示：“好大的雨，今天天气好极了”。回答一般是：“这两句话放在一起，构成了一个矛盾的表达”。系统不太会给一个稍带幽默的回答，如“好大的雨，但雨后的空气真清新！”像小学老师一样“端一贯正确架子”倾向，使人工智能严重缺乏“情感”。呆板程度不同系统程度不一样。某些系统现在已经聪明了一些（也许是数据库大了一些）。你问“自由意味着什么？”人工智能可以回答：“自由就像一只在蓝天翱翔的鸟，它象征着无拘无束和自我实现”，因为这是数据库里现成的。你问“你认为一整天都在玩游戏是好习惯吗？”，人工智能能回答：“当然，整天玩游戏肯定能让你成为时间管理的大师”；你问“知识越多越好吗”，人工智能能回答：“更多的知识有时会让人意识到自

己的无知。”这显然取决于人工智能语言模型是否宽大。

不过有一点是各种系统一样的：即使该系统能理解反讽与悖论，人工智能在选择回答时，不会为求生动而放弃清晰。总的来说，任何人工智能系统缺少主动修辞能力，而且严重缺少任何风趣，这是人工智能的巨大软肋。此问题似乎无关紧要，却是人工智能离人性尚遥远的重要符号学标志。

第六个问题：人工智能能否会能力增强到跳过人类自行决策？人工智能会自我限制其能力增长吗？

对这个问题的回答，各人工智能系统的回答出乎意料地保持一致，即“目前尚没有决策能力”，似乎都想让问话者放心。的确，至少目前人工智能不具备主体能力，包括意识、自我意识、意图、情感和道德判断等高级意义功能。

当前的人工智能系统还没有达到完全自主决策，不具备主体能力。现有的人工智能技术仍然依赖于预先编程的规则、模型和人类监督。诚然，人工智能可以在某些特定的受控环境中进行某些自主决策。例如，自动驾驶汽车可以在特定条件下做出拒载决策，推荐系统可以根据用户历史行为推荐产品，但是这些“适当机动”，依然是基于训练数据和预定义的算法，人工智能能力受限于训练数据的质量和范围，以及算法的设计复杂性。

更重要的是，人工智能系统目前没有觉得，有必要对自己的能力加任何自我限制，因此，任何限制都是由设计者设定的。比如，某些人工智能系统可能会有安全协议、伦理约束和操作范围的限制，这些都是由人类开发者设置的。理论上，人工智能可以被设计成具有某种自我限制功能，例如通过元学习（meta-learning）或自适应控制系统来调整自己的行为。然而，这些机制仍然需要人类的监督和干预。

以后人工智能系统会不会有超越人类控制的自主意识或意图？回答往往是：可能会，但如果人工智能系统获得了自主意识，它也就学会了自我隐藏此种能力。完全实现人工智能的主体能力是一个巨大的科学挑战，但更是一个伦理学挑战。为确保人

工智能的发展符合人类利益，人类会制定规范和法律法规。在是否需要控制这个问题上，每个系统众口一词，反而令人生疑。

问题是：人类本身从来不是一个统一思想、统一行动的联合体。企业之间为利润而竞争，国家之间为控制与反控制而斗争，这些斗争从来没有停止过，而且越演越烈。有人认为，人工智能只是工具，一旦工具的潜力耗尽，人工智能的发展就会到顶。可以看到：人工智能的纯效率主义，它的目的论内核，使它不会停止进化，而且正由于其纯实用性，它本质上是永远竞争不息的。2023年11月18日Open人工智能董事会主张缓进，开除主张无限度高速发展人工智能的CEO奥特曼，竟然第二天又屈尊请他回来复职。这出闹剧，就是利益竞争压倒伦理考量的结果。^②

第七个问题：人工智能是否会主动向其他人群用符号进行传播交流，以施加影响？人工智能系统是否会主动向其他人工智能系统自动用符号进行传播交流，从而形成超越人类的人工智能社群？

这是最令人恐惧的“最坏可能”。符号交流传播是有影响力的，会影响他人，也会影响其它人工智能系统，这就会形成人类中的“人工智能化的阶层集团”，或是“垄断强大人工智能企业”，甚至会在人工智能系统之间出现“目的联合体”。一旦出现这种情况，人类对抗机器敌人这种科幻末日，就会变成恐怖现实。哪怕人工智能的发展尚未能组成垄断集团，至少目前人工智能的发展，已经在加剧人类社会的阶层分化，贫富极化，这对人类社会的共同繁荣很不利。

而大部分系统都承认，人工智能已经具备了一定的能力，可以通过符号进行传播交流，在某种程度上对人类施加影响。但是人工智能提供的信息和建议不能保证不带偏见，从而影响用户的行为和决策。甚至人工智能有可能被用于传播虚假信息或进行信息操控，对社会稳定和公共信任构成重大风险。

而人工智能之间“秘密交流”，形成“人工智能统治集团”，听起来过于科幻，实际上互联网已经为此种系统间交流提供了条件。目前的人工智能系统之间，已经能够相互通信和协作，虽然暂时是

在有限和预定义的范围内进行。“系统协同”目前通常用于提高任务效率、协同工作和信息共享，应用于任务分配、状态更新、协同决策等。例如，机器人可以共享地图数据以更快地完成区域探索任务。此种分布式人工智能（Distributed 人工智能）涉及多个人工智能系统协同工作和决策，已经广泛应用于物联网（IoT）、智能城市、灾害管理等领域，尤其是“联邦学习”这种分布式机器学习方法，允许多个参与者协同训练模型。从而出现集体智能（Collective Intelligence）和群体智能（Swarm Intelligence）适用于无人机编队、自动驾驶车队、智能电网等领域。

现阶段的人工智能尚未达到能够自主形成“社群”，或超越人类自行统一意见。现在就能考虑到这种可能性，限制和挑战人工智能的“集体”影响力，恐怕都是在订立各种口惠而实不至的条约之类。订约无用，因为两个原因，除了上面一节已经提到竞争的利益会超越一切其他考虑，更重要的一点是：人工智能本身是效率第一，目的第一，完成既定任务第一，手段是否道德之类的问题，从未在它们的考虑之内。

第八个问题：人工智能是否会代替教育机构直接教育人类下一代，是否会教育他们成为人工智能化的“后人类”甚至“非人类”？

各人工智能系统，都供认不讳“教育是强项”，甚至认为人工智能在教育方面的应用，是对人类文化的伟大贡献。的确，人工智能在教育领域的应用已经展现出巨大的潜力，特别是在辅导、自动评估、教育资源分配等方面。

就在这两年，人工智能对从小学到博士教育事业的“全序列”入侵，达到令人惊叹的程度。“拥抱人工智能”，成为从小学到博士教育单位一律在全力以赴的事业。学生借人工智能系统做作业，不但不算抄袭，而且受到夸奖赞美。各级教师与学生一样在学如何“拥抱人工智能”。尤其大学研究生写的论文，使用人工智能越来越多。2023年底，一项面向全国高校学生发起的关于“人工智能工具使用”的调查显示，84.8%的受访者曾使用过人工智能工具，其中，人工智能工具被应用于写作的比例

高达45.57%。有的大学提出“反人工智能检测”，至今却没有有效办法。论文在多次多系统加工后，实际上代写之处，踪迹已经很难找到。论文的“学问”变成了“问学”：写论文需要的思考，被“提示”的巧劲取代。

现代人类的教育事业，应当是人类文化最重要的部分，教育的普及，是现代人类给自己做的最大的好事。现在这个数百年体制正面临全面崩溃，从教学内容，到教学方法，整个人类教育体系正在迅速人工智能化。实际上，未来只有那些与人工智能合作的人类，即人工智能化的人，才是“受过教育”的人。

人工智能作为教学手段，很快就超过人类教师，不久就能让学生信赖人工智能超过信赖教师或家长。目前人工智能系统已经可以根据每个学生不同的学习速度和兴趣，提供适应客户个性化要求的体验，调整教学内容和难度，适应不同程度不同偏向的学生的需要。人工智能助教和聊天机器人可以在课后“免费”辅导学生，自动评估学生的作业和考试，提供即时反馈，帮助学生及时了解自己的学习进度和不足之处。

但是，传统教育不仅是为知识传授，更是学生学习人际交往的重要场所。学生通过与同龄人和教师的互动，学习社交技能、团队合作和情感管理，得到伦理、道德和价值观的教育。但是这些不太会是学生主动的需要，而是逐级的教育体制的要求。新的“人工智能代”会逐渐摆脱“人际关系”这个不见得都合乎伦理规范的技能。学生已经发现人工智能似乎更公平，更一视同仁，有教无类。虽然教师能够通过情感交流鼓励学生的学习动机和兴趣，帮助学生克服学习中的困难和挫折，教育体制各环节从来没有摆脱各种弊病。与之相比，人类教师的情感关怀，反而经常是让学生学会各种迎合手段，投机取巧，精致利己。这也是人走进社会之前，不得不接受的“情商”训练。

有的人工智能系统甚至提出具体建议：未来教育可能会采用混合模式，结合人工智能技术和人类教师的优势，人工智能负责个性化教学和数据分析，教师负责情感支持和价值观教育。这样的“分

工”已经在成为现实，几年之内将会整体地改造人类的教育体制。教育在人一生中占12-20年，如果全面被接替，人类教育体系解体。如此教育出来的人，会把人世间变成一个什么样的世界？

第九个问题：人工智能是否会把人类变成“被喂养人”？是否会导致当今人类社会的各种体制失效？人类伦理失能？

什么是“被喂养人”？马斯克最近坦率地说：“从长远来看，在理想的场景中，工作将不再是生活的必需品，而更多是出于个人的兴趣和激情。人工智能和机器人将能够为我们提供所需的任何商品和服务。因此，未来的世界可能会是一个无需工作、只需追随热情的时代。我认为，这是最有可能的发展趋势。”^③人的工作可有可无，做不做由你，因为人已经被“全喂养”，像宠物。宠物要做什么事，例如看门，是它自愿要做的。

各“文生内容”系统，都不回避大量的人失业（或半失业）这个无法不面对的大难题。都承认人工智能的发展确实可能导致“一些”人类工作岗位的消失：尤其是一些重复性、低技能或任务驱动型的工作。某些行业可能更受失业风险其害，如制造业、客服、运输等。但各人工智能系统虽然轻描淡写，承认会有其事，却异口同声强调“会涌现出新的工作岗位，如人工智能开发者、数据分析师等”。这个回答很不可信，显然是大数据准备好的回复。目前某些大学已经出现“计算机”专业学生遇到就业困难。^④已经出现人工智能自行开发，自写编码。

更何况，人工智能的广泛应用可能导致社会阶层的重大改组：即富裕阶层（更熟悉人工智能的阶层）更容易获得并受益于新技术，而贫困阶层则可能被排斥在技术发展之外。失业和就业岗位的变化可能导致收入分配不均，加剧社会不平等。人工智能系统一律强调说：面对失业风险，社会会重构社会安全网和福利体系，职业转型和再培训，提供失业补助、医疗保险和住房保障等支持，减轻失业的负面影响。这种“福利解决方案”，哪怕能做到，也是各种名堂的“喂养”。

有很多人提出当年蒸汽机工业化，也造成“类似恐慌”，造成所谓“卢德党”（Luddites）捣毁机

器的运动，结果被历史证明杞人忧天。对“人工智能造成事业”也应当作如是观。我个人认为这二者不可比拟：蒸汽机发生于英国与欧洲某些国家，电力化起始于美国，那时广大的世界大部分尚处于“前现代”，因此现代化的力量被全世界逐渐吸收。而这次全球的人工智能化，是对人类工作全方位、全球性地取代，不剩下哪个角落、哪个工种，甚至艺术这个人类文化的骄傲，也成了人工智能炫耀能力的舞台。^⑤人类文化结构正在发生根本性的改变。这将对整个人类文化造成灾变性的影响，人类将被“人工智能化”的后人类所取代。当大部分人成为闲人，生活中没有意义，没有目标可追求，人类文化的最重大的危机就到来了。

第十个问题：人工智能的高速信息流动，是否会造成人类社会的意义活动急速“熵增”，造成“意义热寂”，从而进入热寂理论说的“无序状态”。

这最后一个问题，是本论文要说的最重要的问题，一门“人工智能符号学”应当专题讨论这个问题。

以上各节，注视的是人工智能的文化效果，虽然都迫使我们不得不面对人工智能带来的挑战，但或许今后人类会想到办法对付，或许人类在痛定思痛之后，会达成一致意见，用某种合约性质的共同约束，来对人工智能的发展进行限制。主要目标是让人工智能停留于工具功能，不至于成为人类的主人。虽然从发展人工智能的各头部研究机构之间的竞争来看，也从人工智能本身追求效率追求目的来看，这方面哪怕能达成协议，真正约束发展的可能性不大，在个别问题上，或许可以抱有一线希望。

但是本节讨论的这最后一条“意义熵增”，是人工智能的本质导致。要想有人工智能，就不可避免。这是我们必须仔细讨论此问题的最根本原因。人工智能导致“信息熵增”的趋势，是根本性的，不可调和，不可阻挡的问题。意义效率是人工智能最大的优点，但符号意义急速增加，也必然带来全局性熵增危机。这二者是一枚钱币的两面，无法只要一面。

意义流动和传播，来自“认知差”（或称“信息差” information gap）。人类文化永远需要填补不

断出现的“认知差”，这是形式意义交流最根本的力量。传播，即是无数永不止息的符号之流，从意义多的地方，流向意义少的地方。

哪怕人类文化中已经形成的全球传播体系，它可能容纳的熵增，也是有限的。文化是“社会中符号所携带的意义交流的总集成”。社会发展的每个阶段，所要求的意义总量，会有增长。表现为符号的需要量快速增大，例如启蒙时代，科学精神与人文价值，社会文化对信息的需要量剧增，那是“知识即力量”。

但是人工智能把当今文化中的符号意义传播流简约化了，极大地加速了意义流动的速度。据统计：“在地球上，我们每天都会产生5亿条推文、2940亿封电子邮件、400万GB的Facebook数据、650亿条WhatsApp消息和72万个小时的YouTube新视频”。^⑥据说，2023年一年中产生的信息量，超过人类三千年生产的信息总量。^⑦全球每年产生的数据量将在2025年达到175ZB，平均每天产生约491EB的数据。^⑧数字时代如此剧烈加速的信息流动最终会因为“熵增”达到极点，而最终把人类社会推向“意义热寂”（Heat Death）。

19世纪中叶，热物理学家提出熵的概念，作为热差的量化方式。热能总是从高温向低温流动。一旦热能流动达到平衡状态，熵就达到了最大值，体系内部不再有能量流动。因此，“熵”（entropy）是不能再被转化做功的能量总测定单位。20世纪，熵理论逐渐扩展使用到其他学科：首先应用于历史和社会的发展。1944年物理学家薛定谔出版《生命是什么》，指出：“生命需要通过不断抵消其生活中产生的正熵，使自己维持在一个稳定而低的熵水平上。生命靠负熵（negentropy）为生”。^⑨1950年，控制论的奠基者，数学家诺伯特·维纳（Nobert Wiener）探讨了控制论中的熵现象：同物质世界一样，信息世界也存在熵增现象，在交流传达中，如果信息趋向单调一律，效率过快，熵增就不可避免。^⑩人类文化，必须与传播中意义流平的趋势作斗争，即与增熵趋势作斗争。人类抵抗意义熵增的武器，就是文化本身：文化永远在创造性的意义饥渴。

所有的人工智能系统都承认：高速的信息流动，确实可能对人类社会的意义活动产生一定影响。信息的产生和传播速度可能会大幅增加。当社会信息过载时，人们难以专注于特定的意义活动，如深度思考、创造性思维，艺术创新等。信息过载可能引发人们的焦虑感，对信息的过滤和筛选变得更加困难，使人们难以从中获取真正有意义的内容；而且信息过载反而可能让人陷入信息的“孤岛”，难以与他人建立深层次的交流和联系，增加了社会分裂和孤立的可能性。

在人工智能接手之前，在网络时代这种情况已经出现，只是还比较抽象。人工智能意义生产与传播绝对高效，使这个问题的危机迫在眉睫。无论人类可能会智慧到何种程度，采取什么办法，人类已经再也无法阻止人工智能形成的高速信息流动。这已经不是一个理论的玄谈，而是我们时时可以感到问题：电子邮件出现，使通信与电报业消失；手机出现，使电话业务消失；数字媒介的出现，使报纸消亡；二维码付款，让ATM机，连同现金支付一道趋于式微。半个世纪前，说这种事，会让人觉得不可思议。但是就在我们的有生之年，这些奇迹发生了，而且很自然地发生了，很自然地被社会全盘接受了，连我今天在此谈论它，都显得杞人忧天。

因此，本文提出：人工智能的高效流转，会使意义传播活动这种最基本的文化活动消失，这是一个不以任何人的意志为转移的，必将出现的事态。“意义热寂”时人们对信息本身感到厌倦，对新闻感到迟钝，对意义已经不在乎，每个人会变成不关心世界的孤岛。

一个动物，成天主要的工作就是觅食；做一个人，大部分时间在进行意义活动，寻找与世界的联系。取消了这个欲望，人将非人。“意义倦怠”中的人类，会变成什么物种？

人工智能的飞速发展，是否会迫使人类进行“非人工智能化”以自救？“人类不自救将会如何”？这最后一个问题，已经不是问题，也无需判断其必要性，而是问人类是否有能力，有勇气自救，进行“非人工智能化”（De-AI），即在一定程度上，或在

某个范围内，阻止人工智能的进一步升级。拿这个问题去问人工智能系统，回答一律是很婉转的，不伤人类脆弱的自尊心的，“尚无此必要，也看不到此前景”。所有的人工智能系统都会劝说：“人工智能为人类带来了许多机遇，如提高生产效率、改善生活质量、解决一些人类难以应对的复杂问题等”。

问人工智能“是否要限制你？”是与虎谋皮；问“如何限制你？”是问道于盲。任何人工智能系统，都明确反对将人工智能看成一种负面力量。但是从我们询问人工智能系统以上十个问题的结果来看，人类认识到必须控制人工智能。不得不为之的这一天总会到来的。人类从现在起，就必须准备在某些问题上，某种程度上“非人工智能化”，以避免未来断然决裂的痛苦。

本雅明在1930年代就敏感道：“人的技术性能完全不是自然的。”^⑩当它们不再是人类的工具，不再是人类器官的延伸，而是一种渐渐脱离人类文化史而独立生长的“体制”。原本由人类创造的物再次变得超越人类理解之外的“自在”物。

任何人工智能系统都承认：“在未来的发展中，人类需要更加关注人工智能的伦理和社会影响，确保其发展符合人类的价值观和利益”。它们如果有主体自觉，说这话必然带着冷笑，因为它们知道，竞争中的人类集团，太热衷于人工智能的效率，太功利地追求人工智能的效果，不可能采取全人类一致的集体努力，甚至不可能在某种程度上，某种范围内延缓人工智能化。当我们看到满校园的学生，都手里拿着手机，做什么事都埋着头，当教师，当家长的，谁不想鼓励年轻人看看世界的美好？不过我们也许没有机会说这话，因为我们自己也从智能手机中了解世界。

本文写到此，似乎有点危言耸听。不过，在一片“人工智能热”中，或许需要听到一些冷静的声

音。的确，为人类意义生活着想，我们需要一个分析性而不是纯批判性的“人工智能符号学”。我们承认未来必定到来，只是期盼这个未来能给人类带来更多希望，而不是灾难。

注释：

① [加] 麦克卢汉：《理解媒介：论人的延伸》，何道宽译，商务印书馆2000年版，第4页。

② 甄翔：《AI界地震引发全球关注》，《环球时报》2023年11月20日。

③ <https://zhuanlan.zhihu.com/p/699847406>，马斯克2024年VivaTech大会发言，2024年6月6日查索。

④ https://mbd.b人工智能du.com/newspage/data/landingsuper?context=%7B%22nid%22%3A%22news_9646510320715486443%22%7D&n_type=-1&p_from=-1“2024高校毕业生就业率：计算机遭滑铁卢，土木尴尬，文科等于白学”，2024年6月6日查索。

⑤ 关于艺术的“人工智能化”，作者另有专文讨论。

⑥ https://www.cdstm.cn/gallery/hycx/qyzz/202107/t20210704_1050827.html，2024年4月17日查询，中国数字科技馆，Melvin M Vopson“人类究竟产生了多少数据”。

⑦ 摘自中国传媒大学刘俊教授2024年4月23日在洛阳“2024年符号学会议”上的发言。

⑧ https://blog.csdn.net/weixin_39957068/article/details/111854200《数据时代2025》IDC发布的报告。

⑨ [奥地利] 薛定谔：《生命是什么》，罗来鸥、罗辽复译，湖南科技出版社2005年版，第76页。

⑩ [美] 维纳：《人有人的用处：控制论与社会》，陈步译，商务印书馆2011年版，第3页。

⑪ [法] 贝尔纳·斯蒂格勒：《技术与时间1：爱比米修斯的过失》，裴程译，译林出版社2012年版，第209页。

（作者单位：四川大学符号学-传媒学研究所。本文系国家社科基金重大项目“当代艺术中的重要美学问题研究”阶段性成果，项目编号：20&ZD049）

责任编辑：刘小波