

算法治理的技术迷思与行动选择

北京大学中国社会与发展研究中心主任 邱泽奇

【摘要】 算法对人类社会生活影响的普遍性和深刻性让算法成为堪比自然环境的人工环境，算法影响的利弊两面性以及算法侵害的不断出现将人类推入算法治理时代。伴随算法影响不断扩大的是算法自身复杂性的增强，算法已经从古老的计算演化为实时关系模型的自我迭代与演进，进而让算法的过程治理既在技术上不可行，也在管理上代价高昂。中美欧在算法治理领域的探索实践呈现不同格局，美国从防范算法侵害入手，形成了政府和第三方的问责模式；欧盟从数据保护入手，逐渐与美国的问责模式汇流；中国从网络安全入手实施治理，但尚未形成有法理逻辑和明确操作路径的算法治理模式，迈向问责模式似乎是中国算法治理可能的且有效的行动选择。

【关键词】 算法治理 算法技术 问责模式

【中图分类号】 D922.16 **【文献标识码】** A

【DOI】 10.16619/j.cnki.rmltxsqy.2022.10.003

对算法治理的讨论尽管早在20世纪80年代已初现端倪，可真正广泛进入人们的视线还是始于近些年中美欧相关法律的落地和文献数量的爆发式增长。

1984年是桌面计算机进入流行的初始阶段，也是人们对算法开始警觉的起点。次年，美国国会技术评估办公室的一份报告指出，在过去20年间，联邦政府四分之一的部门已经将计算机系统用于执法调查或情报搜集与分析。^[1]其中的“计算机系统”便内含了算法。随后，同一间办公室的报告《电子记录系统与个人隐私》则明确意识到了算法治理的挑战。报告指出，运用电子记录进行个人画像意味着求解个体特征和行为模式以及其与特定行为的关联，进而直接涉及使用个人记录画像、产生算法歧视等隐私和自由问题。^[2]

其实，早在20世纪60年代后期，人们已有对自动化的担忧，类似于当下人们对算法的顾虑。值得记录的一项反思是对人的特殊性和目的性的强调，认为自动化的设计、制造、存在都是应该围绕人展开的，人的尊严是自动化的第一理念。^[3]此后，人们对信息技术的迭代演进、信息技术的数字化，以及数字技术向生产和生活各领域不断渗透带来的影响一直保持着关注。只是这些关注始终呈现为有限的探

邱泽奇，北京大学博雅特聘教授、中国社会与发展研究中心主任、数字治理研究中心主任、社会学系教授，教育部“长江学者”特聘教授。研究方向为数字社会、数字治理。主要著作有《边区企业的发展历程：国有大企业和企业型城市的集体企业》《中国大陆社会分层状况的变迁》《社会学是什么？》《费孝通与江村》《朋友在先——中国对乌干达卫生发展援助案例研究》《技术与组织：学科脉络与文献》等。



讨，没有形成持续的努力。

人工智能的新一轮爆发^[4]引发了近些年对算法治理的关注。大多数涉及算法治理的文献甚至不谋而合地形成了一种套路式表述，如：随着大数据、云计算、人工智能技术应用的普及，算法已广泛且深度地进入平台经济、智慧城乡、政府治理等政治、市场、社会领域，成为影响人类生产与生活的关键力量，算法在带来数字经济繁荣、政府治理有效、社会发展活跃的同时，也潜藏着政治、经济、社会、法律等一系列风险，且相关风险在逐渐显现。为此，政府或社会亟需关注……。中国知网的数据显示，以算法治理为主题的社会科学文献在2013年左右只有零星几篇，2018年开始呈现急剧上升趋势，2022年前4个月发表的文献数量几乎与2020年整年的数量相若。

在既有的探讨中，尽管不少文献直奔算法治理主题，尽抒胸臆地表达对算法治理的关注，却无意或有意地忽视了一些基本问题，如，算法是一个旧已有之的存在，如今为什么要特别关注？到底什么是算法？如果关注，又能关注什么？本文试图从算法对人类社会生产和生活的广泛和深刻影响的程度入手，阐述算法已经成为人类生产和生活的普遍环境，不得不予以关注；本文还将进一步分析算法的技术逻辑，说明关注过程治理、试图对算法进行实时监测难以实现算法治理目标；最后本文将探讨算法治理的可能路径及中美欧的行动选择。

进入算法治理时代

在套路化阐述算法治理的文献中，人们认为，关注算法的必要性来自于大数据、云计算、人工智能等数字技术的广泛应用。问题是，为何相关技术广泛应用之后必需治理算法呢？换个问法，数字技术的广泛应用与算法有什么关系？不用算法，可行吗？即使运用算法，不治理，行吗？我们认为，回答这些问题是理解算法治理必

要性的前提。

算法古已有之。在中国之外，算法的最早证据可以追溯到古代美索不达米亚的巴比伦数学，在伊拉克巴格达附近发现的一块公元前2500年左右苏美尔时期的粘土碑上的除法算法便是证据，^[5]古巴比伦的天文学采用算法程序计算重大天文事件的时间和地点。^[6]在中国，算法的最早证据则可追溯到《黄帝历》，时间大约在公元前2700年，人们用算法理解天体之间的运动规律，安排人类的生产和生活。^[7]如此可知，算法在人类历史中始终存在，在近代社会也没有停止演化，且早已运用于生产和生活之中。^[8]可是，为什么直到现在，人们才那么关注算法治理？

归纳既有文献，我们观察到三种主要理由。一种理由认为，算法侵害^[9]是触发人们关注算法治理的根由。人们认为，现实生产和生活中已经出现了算法歧视、算法偏见、算法共谋、算法垄断、算法黑箱、算法遮蔽、算法短视、算法霸权、算法操纵、算法劫持、算法剥削等与算法关联的社会、经济、甚至政治侵害，给人类社会的政治、经济、社会生活带来了负面影响。

另一种理由认为，算法风险^[10]是触发算法治理的根由。人们认为，在个体层次，算法强化着信息茧房，带来人的认知窄化风险。在市场层次，算法遵循商业逻辑，酝酿着监控资本主义的风险。在国家层次，算法隐藏着被特定利益集团用于社会控制和政治权力再生产的政治风险。从个体层次到国家层次的风险给人类社会的政治、经济、社会生活带来巨大的不确定性。

除了算法带来的侵害和风险，还有人将算法影响上升到制度层次，认为算法权力、算法规则是触发规则竞争、权力竞争甚至规则垄断、权力垄断的根由。值得特别注意的是，这也是西文文献探讨算法治理的主流。早在中文文献关注算法之前，西文文献就以著作形态探讨算法权力（power of algorithms）了。奥赛罗和佩特里希编辑的论文集^[11]运用历史与现实视角刻画算法对日

常生活的影响，如导航、搜索、保密通信、油盐酱醋、观看电视、导购、游戏等，算法已经介入人们的日常生活，成为影响甚至支配人们日常生活的权力。在一般权力之外，社会学家似乎偏爱算法的社会权力（social power of algorithms）。人们认为，算法正在从人类手里获得决策权，在越来越多的领域或行动中成为人类的代理人，不仅在日常生活中代理个体，还在组织行动中代理组织或机构。如果个体和组织层次的代理行动涉及社会秩序的建构和预测，在逻辑上，算法已经进入政治生活领域，^[12]直接接触到了政治权力。在中文文献中，算法权力也在成为一个重要论题。^[13]

同样，西文文献也早在以著作形态探讨算法规则，这些甚至早于对算法权力的探讨。^[14]莱辛格直言，代码即法律。在短短五年内对著作的两次修订中，莱辛格明确指出，代码建构了数字空间的规则，同样约束着人类行动。计算机代码与制度具有同样的规制效应。莱辛格担心的是，商业公司的逐利性会让规则朝着有利于盈利的方向偏移，进而影响代码规则在数字空间的公共性。通俗地说，如果我们把法律作为人类公共性的产物，那么，莱辛格担心的是，在数字空间，代码会将法律的公共性引向为资本牟利。进一步的讨论甚至将算法规则提升到更高规则层级，^[15]认为算法给宪法带来了挑战，对人的自由及其他基本权利造成不利影响，更不用说对制度、规制、公共政策等的影响了。

的确，算法向生产和生活的不断渗透对人类社会产生着深刻的影响，小到个体的日常生活决策如食物、衣装、出行等，中到人与社会每个层次的关系如群体、组织、市场、社会、国家等，大到人类的基本权利如自由等。问题是，自算法进入人类生产与生活开始，算法的影响始终存在，人们为什么要在当下关注算法侵害、算法风险、算法权力、算法规则呢？换一种方式提问，假设这些影响只涉及极少数人，而不涉及整个人

类社会，我们还需要如此关注算法么？答案是否定的。即使有人关注也无需将其推送到整个社会面前，正如20世纪初人们对生产流水线的关注、20世纪60年代人们对自动化的关注一样，算法可以成为少数人的兴趣或学术探讨的前沿，而无须成为整个社会关注的焦点。

我们认为，社会关注聚焦于算法治理其实是由两个因素推动的，除了算法影响的广度和深度，还有算法影响的覆盖面。遗憾的是，既有文献极少涉及后者。事实上，不仅因为算法影响了从个体决策、行动到国家宪法等人类社会生活的各个领域，还因为算法影响了世界的大多数社会、大多数人，可以说，整个人类社会的方方面面都处在算法的影响之下，人类参与的政治、经济、社会生活，处处都有算法的影响，一些领域甚至根本是由算法在推动。

那么，算法影响的覆盖面到底有多大呢？其实，接受算法影响是有条件的，不满足相应的条件，算法影响便覆盖不到。要观察算法影响的覆盖面，还需从算法影响发生的条件入手。对算法影响条件的分析也从另一个角度帮助我们理解，算法始终与人类的生产与生活相伴随，长期以来算法影响覆盖的人群却一直非常有限，如同流水线、自动化等在不同层次影响的人群始终是有边界的。

在数字时代，算法影响的第一个条件是接入互联网络。没有互联网络的接入，就没有个体或组织行动与数字网络的数据交换，也不会有任何数字化算法介入，无论是个体还是机构或组织都被排除在算法的影响之外。根据国际电信联盟（ITU）的数据，截至2021年，世界上依然有37%左右^[16]的人口没有接入互联网。根据中国互联网信息中心的数据，截至2021年12月，在中国以人口总体为分母的互联网接入率为73.0%，10.32亿接入用户，其中乡村用户占27.6%，城镇用户占72.4%。如果以接入设施覆盖面观察，以最小行政单元为计量单位，中国的宽带覆盖率为



100%，光纤覆盖率为98%。^[17]综合相关数据可以获得的判断是，尽管世界上依然有超过三分之一的人口没有接入互联网，尚处在算法影响的范围之外，中国却实现了互联网接入的充分供给，如果以家庭为单位计算，则人口整体处于算法影响的覆盖之内。在数字接入的意义上，算法有机会影响到几乎每一位中国人，如果愿意，每一位中国人也有机会接受算法的影响。算法覆盖的无边界性，在历史上是不曾出现的现象。一个直接的例子是，新冠肺炎疫情流行的近三年来，一块方寸大小的健康状况二维码，让每一位生活在中国的人都感受到了算法影响的强烈冲击。

算法影响的第二个条件是数据成为人类生产与生活的环境。如果只是接入网络，也不足以让算法影响触达每一个接入的节点（人、组织、事物等）。与算法一致，网络也始终伴随着人类社会存在。人际网络是基础网络，由此衍生的网络林林总总。可是，网络并没有因为覆盖面的扩大而改变网络的本质属性，网络还是人与人、人与事、人与物、物与事等事物之间的一种连接关系。当然，连接的性质会因为关系属性的差异而发生变化，从血缘网络、地缘网络，衍生出供应网络、生产网络、创新网络、贸易网络、社会网络、政治网络等。在网络上，两个节点之间流动的只要不是不断产生的数据，不管流动的是什么，即使伴随着算法，算法的影响也仅限于有限的节点之间，不足以引发广泛的社会关注。可是，当网络上流动的是关系性数据时，算法影响的溢出点便出现了。

数据对算法影响溢出的推动不在于数据本身，而在于数据的量。在互联网上，有人专门关注数字网络产生的数据量，如“数据从不睡觉（Data Never Sleeps）”9.0版^[18]显示，在数字网络上每分钟产生的20类数据中，任何一类的数据量都是人脑的计算能力无法胜任的，甚至也是个人电脑的处理能力难以胜任的。数据量巨大到远远超出人类加上简单工具的处理能力，在历史上也

是不曾出现的现象。数字网络进入社会化应用以来，人类活动产生和积累的数据量早已超出传统的处理能力，不借助算法，人类根本无力运用数据资源进行生产，难以提高生活质量。数据已经成为接入数字网络的人口之生产和生活环境，巨量的数据和应用是人类依靠其生物属性无力应付和处理的，由此带来的直接后果是：算法成为了人类生产与生活必备的外生性能力。这样我们就可以理解，为什么智能手机在日常生活中会成为人们的第一陪伴。

算法影响的第三个条件是算法成为人类生产和生活的必备工具。数据和算法的组合正在从根本上改变人类的生产，推动着人类从工业经济迈向数字经济；也在从根本上改变人类的生活，推动着人类从工业社会迈向数字社会，数据和算法因此成为人类迈向数字时代的标志性指标。

在经济领域，上市公司的市值排名不只标示企业价值规模，也指示关键技术的影响力和社会的取舍。根据公开数据，在市值排名前十的公司中，1990年有6家银行、1家通信硬件公司；2020年则有7家互联网公司、2家金融公司。在过去四十年，以设施设备等硬件为基础，以组织和社会应用为目标，数字技术快速迈过其扩散临界点，进入变革的起飞阶段。中国信通院的数据显示，2020年中国数字经济增加值为39.2万亿元人民币，占GDP的比重跃升至38.6%，^[19]这意味着，中国数字经济已越过技术扩散起飞的临界点而进入起飞初始阶段。预计，2025年中国数字经济占GDP的比重将超过50%，^[20]中国将真正进入数字经济时代。

在社会领域，数字终端已经成为人们的必备工具，且不细说数字企业、数字政府如何运用算法改善生产、改善治理，只从人们日常生活对数字工具的应用中便能看到由数据量推动的算法带来的影响。在中国互联网用户群体中，90%的人使用即时通讯、80%的人使用政务服务、70%的人使用数字支付，哪一项应用的实现都依赖算

法。例如，运筹起点到终点的行车路线是出租车司机的基本职业能力，现如今，司机们却说，离开了即时导航都无法行车了。

在政务领域，省部级政务接入国家政务平台的比例快速上升，居民跨行政区办理业务的便利性极大增强。在政务平台上，实名注册个人数达7.74亿，法人数量达7.27亿。审批等行政许可事项实现网上受理和最多跑一次的比例达82.13%；一半以上行政许可事项平均时限压缩超过40%；政府事项网上可办率超过了90%。^[21]数字连接正在推动中国治理体系和治理能力的现代化。

任何一项终端工具应用的背后都有算法，算法的工具化无处不在。我们甚至可以认为：无算法，不生产；无算法，不生活。在生产领域，数据已成为人类的新生产资料，算法则成为人类新的生产工具。在生活领域，数据已成为人类提升生活质量的素材，算法则成为人类新的生活工具。

人类正在加速进入算法时代，人们之所以关注算法侵害、算法风险、算法权力、算法规则，不是因为侵害、风险、权力、规则本身，而是因为侵害、风险、权力、规则直接影响到大多数人的利益和福祉，直接影响到经济发展的可持续性，直接影响到社会发展的平等性，直接影响到政治发展的公平性。换句话说，在人类迈向数字时代的同时，我们无法回避算法给人类带来的如此深重且本质性的影响。问题是，算法又是何方神圣，何以有如此魔力？在既有的讨论中，一些重要的观点可能源自对算法的误解，为此，下文有必要交代我们对算法的理解。

算法治理的技术迷思

尽管算法早已存在于人类社会，可作为概念，算法的词源却来自波斯数学家花刺子模（Muhammad ibn Mūsā al-Khwārizmī）的名字。花刺子模名字的拉丁语写法（Algoritmi）便是今天人们见到的算法作为术语的拼写原型。算法的

现代意义源自公元9世纪花刺子模向西方译介印度计算方法。^[22]可是，直到计算机和计算机科学进入人们的生产和生活之前，人们都没有算法的概念，正如计算机科学创始人诺斯（Donald Knuth）所说，直到20世纪70年代，“绝大多数受过教育的人都不知道算法，并且说实话，无论如何，也几乎没有必要知道”。^[23]

那么，算法到底是什么？在不同的历史阶段、不同的应用场景中，算法有着不同所指，从最简单的算法到最复杂的算法。可在本质上，算法只是一种求解逻辑。举例尽管常常不是准确的定义方法，却是帮助人们理解概念准确定义最有效的方法，还是举一个例子吧。

在一座超大城市，假设您是乘客，在地点A上了一辆出租车，告诉司机您要去地点B，您希望尽快到达目的地。一位出色的出租车司机通常对服务范围的道路非常熟悉。在您告诉司机目的地的瞬间，司机的头脑中会快速盘算行车路线，譬如，路线一可能会遇到堵车，路线二不是最短路径，路线三可能正好适中，或者还有更优的满足尽快到达目的地需求的路线，只是司机不了解。

其中，从A点到B点，是需要解决的问题。有多个与问题相关的因素，至少涉及交通工具、道路状况、交通动态、选择偏好。给定乘客的选择偏好：尽快；再给定交通工具：出租车；司机的路线选择至少还涉及道路状况和交通动态。再假设道路状况已知，便有了司机的三条路线选择。

如果说司机的三条路线选择是从A点到B点问题的数学解，算法则是司机路线选择的求解过程。在过程中，从A点到B点的路线用时是因变量，交通工具、道路状况、交通动态、选择偏好等是自变量，其中，每一个自变量对因变量的影响并非等价，如交通工具是条件变量，道路状况也可以是条件变量，交通工具与道路状况一起又会衍生出其他变量如路线备选集，交通动态可能



是直接的自变量，却是只能凭借司机经验的自变量。运用与事实相吻合的关联逻辑，建立自变量与自变量的关系结构，进一步建构自变量与因变量的关系结构，形成求解的计算逻辑，就是算法。

这个例子表明，第一，算法是面对问题的，无论是现实问题还是理论问题，总之是需要解决的问题。尽快到达目的地是需要解决的实际问题，如何把某个商品卖给愿意出高价的人是需要解决的理论和现实问题，人们为什么在搜索引擎上搜索感冒药也是需要解决的理论和现实问题。人类需要解决的问题无处不在，故算法从不缺乏需要解决的问题。

第二，算法是通过计算解决问题的。因此，任何用到算法的问题，都需要将问题数学化为可计算模型。其中又内含着一个问题的两面。一方面，要解决的问题可以被数学化即用数学模型刻画，不能建模或暂时不能建模的问题是无法运用算法的，如人类或许暂时还无力将A爱恋B的程度数学化。另一方面，建构的数学模型是可计算的。满足可计算的条件是，至少有用于计算的数据，至少计算可以获得解，即存在着运用数据进行计算而获得的解。没有数据，无法计算；有数据却没有计算解也无法计算。通俗地说，运用计算机不是为了计算，而是为了运用计算机的算力更有效率地获得计算的解。哈雷尔（David Harel）之所以将算法定义为计算机科学的灵魂，^[24]或许理由正是在此。

算法正是每一位出租车司机头脑中交通路线的运筹过程。在这个意义上，算法与人类的存在相伴随。可是，人们关注算法，并不是因为某位出租车司机头脑里的交通路线，一个人的算法并不会对其他人构成影响，除非存在某种组织结构，更不会对大多数人构成影响，进而也不会成为社会关注的焦点。

如果不是一位出租车司机，而是成千上万的出租车和非出租车司机，是几百万辆机动车行驶

在道路上，即同时有几百万个算法在发生，不仅如此，几百万个算法还因为都与道路资源有关而构成相互影响，即每一位司机算法的效率都以其他司机的算法为条件，而且，每一位司机的行车路线、行驶状态、驾驶行为都可以成为影响交通路线选择的因素。在这样的环境下，如果还希望获得最优行驶路线，运用人的生物属性已经无力达成目标，即人类的生物属性没有能力处理哪怕是交通路线选择之类的算法之间的关系。

如果道路条件、每一个路口的交通灯变化、道路上每一辆机动车的行驶路线、行驶状态、每一个驾驶者的驾驶行为都被数据化，都汇集到一个数据池，且希望依据数据来获得最优行车路线，即尽快从A点到达B点；那么，此时出租车司机面对的不再只是交通路线选择，而是多出了许许多多工作，且不说相关数据如何测量、汇集，只说汇集后如何识别、整理与运用，每一项活动都是人类的生物属性能力难以处理的，唯一可以借用的，只有算法。

第三，算法是一系列问题和求解数学模型的集合。从一个人的算法到一座城市的算法，在数量规模扩大的同时，算法也完成了它的华丽转身，从一个数学模型变成了无数个数学模型，且无数个数学模型之间不是独立的，相互之间不仅有关系，而且关系还异常复杂，可能形成有条件的、多层次卷积的、多卷积关联网络的或更加复杂的关系，且模型之间的关系不仅有结构，还有时序，结构与时序之间还不是确定的。以前置算法为条件的算法以及人类对计算效率的追求，还促进了算法的自我迭代，即以算法为数据而改进算法。

算法不仅与算法交互，也与人类行为或事物的发展变化交互。还是以城市道路交通为例，道路上的每一位机动车驾驶员都有随机应变的能力和行动，而每一个随机应变都有机会改变前一个时点算法的结果，也意味着被运用于计算的算法需要面对人类行为的改变、道路状态的改变、交

通环境的改变而迭代和优化。如此，算法的实际运行远比人们想象的复杂，一个看起来只呈现了红橙黄绿交通状态的即时导航系统，却是异常复杂算法融合的结果。以谷歌流感模型为例，算法科学家运用用户搜索数据呈现的2007年流感模型是在4.5亿个模型中筛选综合的结果。^[25]

可是，如何把不同的算法整合到一起形成求解模型的集合呢？且不说4.5亿个模型，即使是45个模型，也是人类的生物属性无法处理的，数45个数还常常出错呢，算45个模型且将45个模型建立关联，更不是人类的生物属性可以无错应对的了。那么，如何让用于解决问题的模型不变成进一步需要解决的问题，而真正解决人类问题呢？算法对代码的需要出现了。

第四，算法是由代码连接且结构化的一系列问题和求解数学模型的集合。通俗地说，任何由计算机执行的指令都是代码。复杂算法的代码量可以多达数亿行，可代码向计算机传递的每一条指令却只有一个动作，做（do）或不做（else），再复杂的数学模型都需要将其转化为向计算机传达的简单指令。复杂的简单也因此变成了复杂的复杂的一部分，计算机代码也因此成为了算法的有机组成部分。实现算法代码化的不是别人，正是人们将数学模型转化为结构化的计算逻辑（如Alonzo Church）和被称为电子计算机鼻祖的图灵（Alan Turing）将计算逻辑转化为可由计算机执行的指令逻辑，其中，每一条指令都以代码形式呈现。在算法代码化的进程中，代码（code）的确是规则，代码的确在指示计算机做或不做。由此推演下去，表面上，正是由代码携带的简单指令构成的复杂算法影响着计算结果的输出，带来了文献关注的算法侵害、算法风险、算法权力、算法规则。可是，一旦我们进一步了解了代码与算法的关系，或许会有不一样的观点。

综上所述，如果仅从技术上观察，算法不过是面对理论和现实问题的数学求解过程，只是因为求解面对的复杂性超出了人类生物属性的能力

范围，人们创造出一种机器，且机器能够理解人类将复杂求解过程简化为做或不做的机器指令，交给机器去求解而已。通俗地说，理解当下热议的算法治理，我们必须认识到，算法对人类社会的影响是随着第一台可编程电子计算机和第一种计算机编程语言的出现而出现的，随着计算机和计算机科学的发展，算法才真正开始在解决人类问题的需求中发挥核心作用，从初期的军事和科学领域，到商业和管理领域，再到当下的日常生活和生活领域，算法的影响随着算法应用的拓展和对人群覆盖性的扩大而呈现。“是计算机科学的快速发展，特别是对算法的研究，改变了这一切。如今，算法已变得必不可少。”^[26]只是，从解决问题的需求到问题解决的数学求解，从极为简单的数学求解到数学求解之间的关系结构化、复杂化、动态化，算法已经不再简单如出租车司机的行车路线运筹，在数字时代，算法已经转型为一系列关系和关系动态的代码化，问题、模型、代码已经成为算法的内核。

算法治理的行动选择

上文讨论已经证明，算法之所以会成为社会关注的焦点，不在于算法有多复杂，而在于算法会对许多人甚或对整个社会带来影响。如果说数字连接、数据丰盈、人类必备是算法影响的三项必要条件，问题求解对能力的要求超出人类生物属性的极限是算法影响的充分条件；那么，在人类跨入数字时代、中国进入数字发展加速期的当下，算法又如何产生影响呢？我们认为，只有理解算法影响人类的机理，才能从出发点找对算法治理的方向。

既有文献对算法影响社会的机制认识大致可以归结为两个本质性的争议。第一，算法是中立的还是有偏见的？斯廷森（Catherine Stinson）指出，人们以为算法是中立的，干净的数据和良好的意图可以消除偏见，其实，算法并不是中



立的，除了有偏见的数据和算法生产者以外，算法自身也会产生偏见。这种偏见不一定来自于数据，也不一定来自算法生产者，而是来自于机器学习的进程，算法在不断的迭代中从用户对算法推荐的响应中进行学习，创建着针对用户的选择性偏差，即偏见。^[27]这些偏见不仅具有统计学意义，也会带来不同类型不同程度的算法侵害后果，如学者们笔下的算法共谋、算法垄断、算法黑箱、算法遮蔽、算法短视、算法霸权、算法操纵、算法劫持、算法剥削等，即我们统称的算法侵害，而算法决策的普遍性会让算法侵害对人类社会构成侵害。未来的侵害对当下而言即是风险，即学者们说的算法风险。^[28]

第二，算法是人类意志的产物还是自主自为的？认同“算法即规则”^[29]观点的研究者认为，算法已经作为独立行为主体参与人类生产、生活，算法规则也外化为社会行为规则，对人类产生深度影响。例如，算法的自动化和智能化特征可能诱使技术研发者和应用者在提高效率、降低成本的引导下，更多追求控制乃至替代人类行为的自主性。^[30]有人甚至认为，算法的自主自为性形成了算法利维坦，成为建构人类环境的重要乃至决定性的力量，不仅每一位社会成员正逐渐在成为算法的附庸，国家权力运行也在日益依赖算法，社会的技术理性至上实践建构了算法的绝对权威。^[31]

在探讨算法治理行动之前，我们认为有必要针对算法影响机制的争议进行澄清和回应。实践总是最客观的出发点，让我们通过一项实证研究来厘清。一项对谷歌图片算法的研究，对4个以上种族所有年龄段的4万多张面孔的算法展开的审查显示，谷歌图片算法再现了白人男性父权制结构。算法对白人男性有更多样化的和更积极的刻画，对其他人群的描述则更简单，更加呈现如刻板印象、歧视女性等社会歧视。^[32]不过，研究团队的工作没有就此停止，而是借鉴布迪厄的文化再生产理论对算法进行再培训。结果显示，经

过再培训的算法呈现了算法对社会文化实践的刻画，改善了之前呈现的社会歧视。^[33]

这项研究告诉我们，如果代码生产者以客观为由，真实地拟合社会现实，那么，算法的确会携带偏见，进而制造算法侵害。复刻部分人群（如代码编写者、数据汇集者等）的社会歧视和刻板印象，不只是将现实社会的社会歧视和刻板印象算法化，还意味着将部分人群的社会歧视和刻板印象通过高度互联的社会网络传递给不明真相的用户，因此也不再只是布迪厄意义上的文化再生产，而是对社会歧视的扩散与传播，甚至是社会歧视的教唆。

对算法再培训带来的算法改善则清晰地说明，第一，算法侵害看起来是算法偏见制造的，实际却是编写算法的人群制造的，算法只是做了算法可做的事：有效率地运行代码，为人类提供计算结果。无论是通过算法迭代还是自编码，算法都只是放大人类在初始阶段植入的偏见和规则，而不是算法自主自为地形成偏见和规则。以为算法存在独立于算法生产者之外的偏见或规则，以为算法不再是人类生产的产品，其实是一种方向性的误解。第二，算法并不具有自主自为性。算法的确可以自我改善，自己编写代码，可是算法却没有自带偏见、也没有自造规则，即使有偏见、有规则，也是人类将自己的偏见和规则变成算法代码形成的偏见和规则。没有人类输入的第一行代码，便不会有算法的其他代码，算法的自主性自为性便失去了载体，也根本没有机会。第三，支持算法偏见和规则影响力的不是算法本身，而是对算法的运用。如前所述，算法的个体应用不会对其他人群产生影响，算法的平台化运用则是促成算法影响力的机会和力量，无论是算法偏见、算法规则，还是其他。^[34]当然，算法的平台运用并非只是出租车司机头脑里的交通路线运筹，而是更广泛的行动关系（复杂算法、网络算法），一方面是面对问题的技术解决方案，另一方面是平台组织与技术解决方案的关系

体系。同时，由于技术解决方案和平台都是面向用户的，庞大的用户规模又具有社会性，最终形成的是技术解决方案、平台组织、社会之间的不仅关涉生产效率，还牵扯社会平等的综合关系体系。

如此，算法治理的方向选择直接摆在了我们的面前。第一，算法不具有自由意志，算法是人类行动的产出，没有脱离人类行动的算法，从最简单的算法到最复杂的算法，都是如此。因此，算法治理是针对人类生产算法的行动治理。把算法偏见归咎于算法本身，无异于同风车作战。有人依据算法承担的任务将算法在人类行动里承担的角色划分为纯粹执行者、辅助决策者、独立决策者。^[35]这一划分看似来自实践，实际上却对算法自由意志的设定含混不清，一方面假设算法是人类产品，另一方面又假设算法具有自主自为性。

第二，算法治理一方面是针对算法偏见的治理，另一方面也是针对算法平台运用的治理。算法治理的目标，前者是为算法纠偏，后者则是约束算法运用的方向和范围，以期最大限度地减少算法偏见带来的算法侵害，维系和促进社会平等。在现实生活中，算法的主要生产者平台，制造算法影响力的也是平台，是故，算法治理便可以汇聚为针对平台的算法治理。

在中文语境中，算法治理有双重含义。第一，用算法进行治理（governance by algorithm）。第二，对算法进行治理（governance to algorithm）。英文语境却没有算法治理（algorithm governance）概念，也很难找到一个对应的概念，维基百科中甚至没有相应词条。纵观中文文献对算法治理的探讨，大多探讨的是第二重含义。在接下来的讨论中，我们也在第二种意义上使用算法治理，在给定算法不具有自由意志、且不具有自主自为性的前提下，探讨针对算法偏见带来的算法侵害之治理行动。

算法治理发端于美国，^[36]美国如何应对算法

侵害呢？直接的答案是，问责模式，^[37]问责对象是有潜力让算法产生广泛且深刻影响的算法生产者和使用者，问责内容包括算法内容和非算法流程，问责的主体是政府或第三方。

2017年1月12日，美国计算机协会下属的美国公共政策委员会（USACM）发布的《算法透明度和责任声明》指出：“生活中无处不在的算法促使我们应该专注于解决与算法设计和技术相关的挑战，并从一开始就防止偏见。”围绕算法公平，从建立算法公平意识到实现算法公平的算法检验，这项《声明》提出了7项促进算法透明度和可靠性的原则，建立了算法治理的外部问责和控制模式。

在州、市一级的算法治理中，2017年12月，纽约市议会通过了第一部针对人工智能监管的《政府部门自动决策系统法案》（Auto Decision System Task Force Law），要求成立包括自动化决策系统专家和受自动化决策系统影响的公民组织代表在内的工作组，对市政机构运用自动决策算法的公平性、问责性和透明度等进行监督，由此开启了算法治理外部问责与控制的实践模式。

2019年4月，美国国会参众两院通过了《算法问责法案》，授权美国联邦贸易委员会（FTC）要求并监督企业对其人工智能系统的“准确性、公平性、偏见、歧视、隐私和安全性”等问题进行自我审查，且约定这一法案适用于年收入超过5000万美元或拥有超过100万用户数据的数据代理商和企业。由此建立了治理对象的范围。

2021年5月，美国参议院引入《算法正义与在线平台透明度法案》，要求规范算法过程，进而保护个人信息，保障个人特征不因算法而受到歧视。《法案》建立了算法安全性和有效性标准，对歧视的判断标准进行了详细解释，还要求在线平台为非算法流程承担责任，包括算法运行的可靠性、个人数据的可携带性和将个人数据作为司法证据的可实现性，由此建立了将个人数据



证据化、将算法偏见归责化的逻辑闭环。

简单地讲，美国把因算法偏见带来的不良社会后果直接归责到算法生产者和使用者的，而没有归咎于算法自身。

在数字时代的三方国际力量中，欧盟是重要一方。尽管欧盟没有针对算法的专门法律，可针对数据保护的一系列法律在实践中也直接约束了算法的生产和运用。1995年颁布的《数据保护指令》确认了数据主体的权利、数据保护的监管机构，对数据的跨国流通等也作出详细规定，建立了以数据保护为出发点的算法治理的逻辑起点。2000年颁布的《电子商务指令》则规范了数字服务提供者的责任。2015年欧盟进一步出台的《数字单一市场战略》虽不具有法律效应，却在理论上奠定了针对数字市场的监管框架，把算法产出（如搜索结果、广告链接等）透明度，数据流动的组织间关系、数据流动的意愿与约束等都纳入市场监管范围，为算法治理指出了方向。2017年颁布的《机器人民事责任法案》则把算法偏见带来的算法侵害，如对个人决策带来的负面影响，对政府行政、司法、公共政策带来的不利影响等，明确归责于算法的生产者和使用者，并且在道德层面对算法生产和使用提出了要求。由此，我们观察到，欧盟的算法治理从守势的数据保护即通过保护数据来保障权益也转向了不得不面对的算法侵害，并与美国一致，将算法侵害归责到算法的生产者和使用者，而不是算法自身。

欧盟2018年颁布的《通用数据保护条例》（GDPR）将自1995年以来零碎的算法治理理念具体化为数据保护规制的重要组成部分，更具体地定义了数据，对数据的提供、汇集等作了更明确的规定，还详细约定了数据主体的数据访问权、纠正权、被遗忘权、限制处理权、反对权、拒绝权和自主决定权等权利，更对数据的处理和运用进行了严格的限制，明确禁止可能造成算法侵害的数据运用，如揭示种族或民族出身，揭示

政治观点、宗教或哲学信仰，刻画自然人健康、性取向等。欧洲议会秘书处2019年发布的《算法问责和透明治理框架》（A governance framework for algorithmic accountability and transparency）提出了负责任的研究与创新（RRI）概念，从道德、法律到实践，明确了算法治理的问责框架。欧盟委员会2020年提案的《数字服务法案》和《数字市场法案》在数字治理上则进一步向美国靠拢，强化了算法生产者和使用者的责任；2021年提案的《人工智能法案》则区分了规则算法（rule-based algorithms）和学习算法（learning algorithms），建立了基于算法风险的算法治理改进路径，再次强调算法生产者和使用者的责任，强调问责机制。

总体上看，虽然欧盟与美国一样在形成算法治理的逻辑闭环，却在实践中从数据保护入手，建立了保护个体、组织不受算法侵害的防御性治理框架和法律制度，在算法治理理念上则已向美国靠拢，即采用政府和第三方问责模式治理算法可能带来的侵害，且将算法侵害责任归咎于算法生产者和使用者。

到这里为止，其实我们需要问一个问题，为什么美国的算法治理要采用问责模式？为什么欧盟会向美国靠拢也采用问责模式？前一节其实已经回答了这个问题的技术部分，即，算法的复杂性使得算法自身始终处于迭代之中。换句话说，在生产和生活中运行的算法不是一个确定现象，而是一个动态。譬如，人们熟悉且经常使用的乘用车即时导航系统。无论采用哪一家数字平台提供的导航系统，都会有一项服务：道路交通动态实时更新。在从A点驶向B点的进程中，只要出现道路动态变化，系统都会提醒您前方可能的优化备选路线和道路上的事故或拥堵动态，呈现在用户屏幕上的、变化着的、从绿色到深红色的道路标记变化正是实时计算的产出，也意味着算法在被使用进程中的自我迭代。

算法的实时迭代意味着对算法代码进行实

时督导 (monitoring) 和审查 (auditing) 需要具备与算法生产和使用相当或超越的技术能力。仅此, 便意味着在技术上不可行; 或者即使可行, 至少也是艰难的和代价高昂的。对超出能力范围的技术过程进行监管, 需要跳出技术和技术过程本身, 寻求算法治理的关键节点和效率节点, 正如俗语说的, 打蛇要打在七寸上。

那么, 算法治理的关键节点在哪里呢? 算法的生产是一个过程, 从算法设计、代码编写到算法产品; 算法的应用也是一个过程, 从算法部署、算法迭代到算法产出。如果沿着算法的生产与应用逻辑和流程进行治理, 可以称之为过程治理。一旦算法投入应用并达到足够的规模, 算法便会产生相应的社会后果, 包括算法偏见。如果针对算法的产出展开治理, 可以称之为结果治理。算法治理的社会特征在于过程治理的不可行性或代价高昂以及结果治理的简单易行, 是故, 算法治理的关键节点便已不言而喻, 那就是, 针对算法结果进行治理。

算法问责便是对结果的治理, 也显然采用了删繁就简原则。第一, 在法律上约定算法可追责性。要求算法透明、算法解释、数据证据化等, 正是为结果可追责提供事实前提和法理逻辑前提。第二, 在规则上约定算法责任实体性。尽管算法即规则, 可算法规则只适用于计算过程, 不适用于外化的社会过程, 算法生产者 and 使用者才是算法的责任主体, 即算法的责任主体是人或人的集合体, 不是算法自身。一些人以为算法的行动责任已经向机器转移, 事实上并没有, 机器行动依然是人类行动的一部分或人类行动的目的性延伸。如果我们把算法向人工智能甚至通用人工智能领域的拓展理解为人类对效率追求的时代发展, 在提高效率、降低成本的引导下, 即使人类会忽视算法偏见的存在甚至助长算法偏见, 算法也依然只是算法, 算法的自主性是人类赋予甚至纵容的自主性, 在本质上还是人类的行动选择, 而不是机器的主动侵入。让机器为人类的选择偏

见背锅是人类卸责的行动, 不是机器的行动。

有鉴于此, 作为数字时代三方国际力量之一的中国, 又在如何进行算法治理呢? 从1994年制定第一部相关行政指令《计算机信息系统安全保护条例》到目前, 中国共出台了60多部相关法律法规和行政指令。从时间顺序来看, 可以认为中国的算法治理是从总体安全出发的。

虽然早在2000年9月国务院就颁布了《互联网信息服务管理办法》, 可该《办法》更多是行政许可对服务内容相关的约定, 属于数据经营监管类的行政指令。进入算法治理时代以来, 第一部相关法律《中华人民共和国网络安全法》(以下简称《网络安全法》) 虽不直接涉及算法, 却奠定了算法治理的基本方向, 即以安全为依归。2017年6月实施的《网络安全法》的焦点是网络空间安全, 也可以被理解为网络环境安全。网络安全也被广泛传播为国土安全的拓展, 即将网络空间作为国土空间的一部分, 从守护安全出发强调数字基础设施、网络运行、网络信息的安全与保障。第二部相关法律是2019年1月实施的《中华人民共和国电子商务法》, 焦点在于交易安全, 部分内容涉及对算法结果的治理, 如支付指令发生错误时的处理等, 却没有专门涉及算法侵害的内容。2019年10月实施的《儿童个人信息网络保护规定》虽聚焦于个人信息, 却依然强调了信息安全, 这里也出现了一个新的动向, 即对儿童信息相关运营商的责任约定。2020年3月实施的《网络信息内容生态治理规定》依然是安全导向的, 强调网络信息生产与发布、网络内容服务的无害性, 即内容安全, 只有在关于内容服务的第十二条涉及了算法治理, 不过, 依然是从内容安全出发的。第十二条规定: “网络信息内容服务平台采用个性化算法推荐技术推送信息的, 应当设置符合本规定第十条、第十一条要求的推荐模型, 建立健全人工干预和用户自主选择机制。”第十条、第十一条规定的都是内容属性, 而不涉及算法侵害。2021年9月实施的《中华人民共和国



数据安全法》则将网络安全进一步延伸到数据领域，焦点依然是安全，即在总体国家安全观的指导下建立健全数据安全治理体系。

2021年11月实施的《中华人民共和国个人信息保护法》（以下简称《个人信息保护法》）才真正涉及算法治理，出现了与欧盟GDPR相似的内容，不过，安全指向依然清晰明确。个人信息权益、个人信息处理及个人信息跨境流动，都清晰地指向了个人信息安全。值得指出的是，其中也有一个清晰的转向，即在知情同意的意义上，明确了个人信息的个人权属性，区分了一般个人信息和敏感个人信息，说明了个人在个人信息处理中的权利，也规定了违反个人信息处理规则的法律后果，如明确涉及算法的第七十条规定“个人信息处理者违反本法规定处理个人信息，侵害众多个人的合法权益的，人民检察院、法律规定的消费者组织和由国家网信部门确定的组织可以依法向人民法院提起诉讼”。问题是，《个人信息保护法》依然只是原则性规定，还没有完成个人信息保护的司法证据化约定。

2022年3月实施的《互联网信息服务算法推荐管理规定》（以下简称《规定》）才是第一部针对算法的管理规定。《规定》虽然只是四部委的行政指令，却有了针对算法黑箱、算法歧视、信息茧房、舆论操控、不正当竞争等算法侵害的规范。与既往的相关法律和行政指令不同，《规定》明确了信息服务规范、用户权益保护、监督管理责任和法律责任，将监管责任主体明确为四部委（网信办、工信部、公安部、市场监管总局），并将相关违规的法律责任监管纳入国家相关部门管理之中。不过，联系2021年9月国家互联网信息办公室、中央宣传部等九部门联合发布的《关于加强互联网信息服务算法综合治理的指导意见》（以下简称《指导意见》）可以看出，《规定》依然是一部以安全为导向的行政指令，因为《指导意见》中明确指出要坚持正确导向、依法治理、风险防控、权益保障、技术创新等五

项基本原则，要“利用三年左右时间，逐步建立治理机制健全、监管体系完善、算法生态规范的算法安全综合治理格局”的主要目标。即使如此，遗憾的是，到此为止，中国的算法治理只有目标，相关法律法规尚没有像欧美那样形成从治理理念到判例实践的闭环，其中，尤其是还看不出算法治理的法治理理念，更没完成算法治理的司法证据化约定。

如果说算法偏见带来的算法侵害是算法治理面对的核心问题；那么，在给定算法偏见人因化或组织化，给定算法技术处于不断迭代且动态演化的场景下，实施算法治理的行动选择其实不多。

第一，鉴于算法的技术复杂性和动态性，约定算法透明是让算法可以作为算法审查^[38]和追责的前提，却不是实施算法监测的理由。如前所述，算法的技术迷思已经宣告了算法监测的低效率和高成本，因此，过程治理不是算法治理的可选项。可是，针对过程的治理要求却是其他治理选项的必备条件。

第二，如果我们把算法安全风险理解为算法侵害的一个类型，则结果治理即使不是算法治理的唯一选择，也是有效选择。防范风险是保障安全的一种策略，治理侵害则是防范风险的明智之举。在算法无处不在且覆盖广泛而深刻的数字时代，算法已经成为人类生产和生活的环境，这也意味着算法风险时时刻刻无处不在，而不只是学者们列出的个体、组织、国家、政治、经济、社会的风险。对时时刻刻无处不在的风险，优化的行动选择是治理出现的侵害并将其作为未来可能的风险源，而不是治理出现概率极小的、可能出现侵害的预设风险。

第三，算法侵害的人因化意味着治理行动的对象是算法的生产者和使用者，而不是算法。在相当长一段时间内，支持通用人工智能的算法还不会出现，算法依然是人类的产品，即使算法具备学习能力，第一行代码依然是人类编写的。因

此，算法治理的关键是约束算法生产者写对第一行代码，约束算法使用者用对第一行代码，问责模式至少是当下可行且有效的治理模式。

第四，鉴于算法的生产者和使用者以及对人类构成广泛且深刻影响的主要是平台组织，把算法治理纳入平台治理、作为平台治理的一项内容，在形成成熟的算法治理进程中至少可以作为一项过渡性的行动选择。的确，平台组织汇集了人类的代码精英，^[39]用问责模式敦促平台组织约束代码精英在代码生产中融入人类的积极价值观，融入人类共同的伦理，^[40]融入保障人类平等的相关法律和法规，可以从源头治理算法偏见，预防算法侵害，进而最大限度地降低算法风险。

结论

伴随数字连接迈向人类整体连接的是各类数字终端包括移动终端和传感器的时时刻刻无处不在，如此广泛且深刻的数字连接因连接的“节点—中心”结构而形成的是数据的汇流与应用的分发模式。在汇流与分发之间，算法成为运用数据资源促进人类发展和平等的关键所在。

事实是，算法在带来数字红利的同时，也在带来算法侵害。人们关注算法侵害不是因为算法侵害本身，而是因为算法利弊之影响对人群的覆盖。算法古已有之，算法侵害古已有之，之所以没有形成广泛的甚至令人焦虑的关注是因为它不涉及大多数乃至人类社会整体。如今，数字连接已经覆盖了世界人口的67%，覆盖了中国的几乎所有家庭，算法侵害无孔不入，有机会让每一位处在数字网络中的人感同身受，人们关注算法侵害正是源于算法影响的广泛性和深刻性，人们关注算法治理也源于人们试图阻止算法侵害的普遍发生。

如果说算法侵害是算法带来的不利后果，学者们则用更多源于事实的归纳来具体化这样的侵害，如算法歧视、算法偏见、算法共谋、算法

垄断、算法黑箱、算法遮蔽、算法短视、算法霸权、算法操纵、算法劫持、算法剥削等。算法给人们带来的负面影响还可以列出更多，甚至是一个无穷列表。试图阻止算法具体的负面影响并不足以防止算法侵害。

算法侵害源于算法偏见。可是算法偏见不是算法自己的偏见，而是人类偏见在算法中的再现甚至放大。在通用人工智能出现之前（甚至出现之后），尽管算法已经从人类的简单计算活动迈向机器的自编码，算法始终还是人类的产品，而不是独立于人类的、具有自由意志的、与人类并行的客观存在。没有人类输入的第一行代码，便不会有动态演化的机器算法。如此，阻止乃至防止算法偏见还需首先纠正人类的偏见。

针对算法偏见来纠正人类的偏见有两个基本路径，一是针对偏见发生的过程，二是针对偏见带来的后果。落实到算法治理中，前者是过程治理，后者是结果治理。算法的技术迷思决定了过程治理是低效率的和高成本的，算法的动态性也决定了试图通过过程治理来预防算法侵害无异于与风车作战。针对算法偏见带来的结果进行治理在理论上是更加有效的行动选择。

在数字时代的国际三方中，美国的算法治理采用了算法侵害的事后问责模式，同时，也为问责实现如便于算法审查作了算法透明、数据证据化等约定，形成了从算法治理法理到实践的逻辑闭环。欧盟的算法治理从保护数据的守势入手，也在迈向美国的算法问责模式。中国的算法治理以网络安全为起点，以总体安全为目标，尚未呈现清晰的法理，在实践中颁布了一些点位性规则，却还没有形成以法理为依据、以社会平等为依归的算法治理逻辑。我们认为，针对算法偏见的事后问责或许是殊途同归的有效选择。

（本文系国家社科基金重大项目“新技术应用背景下数字社会特征研究”子课题“数字社会的国家治理研究”成果的一部分，项目编号：19ZDA143）



注释

[1] U.S. Congress; Office of Technology Assessment, *Federal Government Information Technology: Electronic Surveillance and Civil Liberties*, U.S. Government Printing Office, 1985, pp. 3-4, 14-15.

[2] U.S. Congress; Office of Technology Assessment, *Federal Government Information Technology: Electronic Record Systems and Individual Privacy*, U.S. Government Printing Office, 1986, pp. 89-95.

[3] Sterling, T. D., "Guidelines for Humanizing Computerized Information Systems: A Report from Stanley House", *Communications of the ACM*, 1974, 17(11), pp. 609-613.

[4] 姜李丹、薛澜：《我国新一代人工智能治理的时代挑战与范式变革》，《公共管理学报》，2022年第2期。

[5] Chabert, J. L., *A History of Algorithms: From the Pebble to the Microchip*, Springer Science & Business Media, 2012, pp. 7-8.

[6] Aaboe, A., *Episodes from the Early History of Astronomy*, New York: Springer, 2001, pp. 40-62.

[7] 陈久金、杨怡：《中国古代的天文与历法》，北京：商务印书馆，1998年。

[8] 邱泽奇：《算法向善选择背后的权衡与博弈》，《人民论坛》，2021年2月中下期。

[9] 肖红军：《算法责任：理论证成、全景画像与治理范式》，《管理世界》，2022年第4期；阳镇、陈劲：《算法治理：成为责任型数智企业》，《清华管理评论》，2021年第4期；丁晓东：《算法与歧视：从美国教育平权案看算法伦理与法律解释》，《中外法学》，2017年第6期。

[10] 孟天广、李珍珍：《治理算法：算法风险的伦理原则及其治理逻辑》，《学术论坛》，2022年第1期；张爱军：《“算法利维

坦”的风险及其规制》，《探索与争鸣》，2021年第1期；贾开：《人工智能与算法治理研究》，《中国行政管理》，2019年第1期。

[11] Ausiello, G. and Petreschi, R., *The Power of Algorithms: Inspiration and Examples in Everyday Life*, Springer Science & Business Media, 2013.

[12] Beer, D., "The Social Power of Algorithms", *Information, Communication & Society*, 2017, 20(1), pp. 1-13.

[13] 张爱军、王首航：《算法：一种新的权力形态》，《治理现代化研究》，2020年第1期。

[14][29] Lessig L., *Code: And Other Laws of Cyberspace (New edition)*, New York: Basic Books, 1999. 第二版于2006年出版，基于第二版的中文版于2018年出版。参见 [美] 劳伦斯·莱斯格：《代码2.0：网络空间中的法律（修订版）》，北京：清华大学出版社，2018年。

[15] Micklitz, H. W.; Pollicino O.; Reichman A., et al., *Constitutional Challenges in The Algorithmic Society*, New York: Cambridge University Press, 2022.

[16] 参见国际电联 (ITU) 网站：<https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx>.

[17] 中国互联网络信息中心：《第49次中国互联网络发展状况统计报告》，2022年。

[18] 参见 <https://www.domo.com/learn/infographic/data-never-sleeps-9>.

[19][21] 主要依据综合自中国信通院：《中国数字经济发展白皮书》，2022年；中国互联网络信息中心：《中国互联网络发展状况统计报告》，2022年。对两类报告均参考了其初期和近两年的数据。

[20] 李稻葵在2021年亚洲博鳌论坛上的发言，参见媒体报道。

[22] Referred to "Al-Khwarizmi biography," www-history.mcs.st-andrews.ac.uk, Archived from the original on August 2, 2019. Retrieved May 3, 2017.

[23][26] Knuth, D., *The Art of Computer*

Programming. Volume 1: Fundamental Algorithms, New York: Addison-Wesley Professional, 1997.

[24]Harel D. and Feldman. Y., *Algorithmics: The Spirit of Computing (3rd ed)*, Harlow: Addison-Wesley, 2004.

[25]Butler, D., "Web Data Predict Flu", *Nature*, 2008, 456 (7220), pp. 287-288.

[27]Stinson C., "Algorithms Are Not Neutral: Bias in Collaborative Filtering", *ArXiv*, 2021. 2105.01031.

[28]孟天广、李珍珍：《治理算法：算法风险的伦理原则及其治理逻辑》，《学术论坛》，2022年第1期。

[30]贾开、薛澜：《第四次工业革命与算法治理的新挑战》，《清华管理评论》，2021年第4期。

[31]张爱军：《“算法利维坦”的风险及其规制》，《探索与争鸣》，2021年第1期。

[32][33]Papakyriakopoulos O., and Mboya A. M., "Beyond Algorithmic Bias: A Socio-Computational Interrogation of the Google Search by Image Algorithm", *ArXiv*, 2021, 2105.12856.

[34]算法不只是因偏见带来侵害，更主要还是因效率带来算法红利，算法红利是数字红利的

核心。人类对数字技术的追逐正是由于数字红利的吸引。参见邱泽奇、张树沁、刘世定等：

《从数字鸿沟到红利差异——互联网资本的视角》，《中国社会科学》，2016年第10期。算法红利不在本文讨论范围，故略去。

[35]肖红军：《算法责任：理论证成、全景画像与治理范式》，《管理世界》，2022年第4期。

[36]许可：《驯服算法：算法治理的历史展开与当代体系》，《华东政法大学学报》，2022年第1期。

[37]孙逸啸、郑浩然：《算法治理的域外经验与中国进路》，《信息安全研究》，2021年第1期。

[38]Shen, H.; DeVos, A.; Eslami, M. et al., "Everyday Algorithm Auditing: Understanding the Power of Everyday Users in Surfacing Harmful Algorithmic Behaviors", *ArXiv*, 2021, 2105.02980.

[39]Burrell, J. and Fourcade, M., "The Society of Algorithms", *Annual Review of Sociology*, 2021, 47(1), pp. 213-237.

[40]Timcke, S., *Algorithms and the End of Politics: How Technology Shapes 21st Century American Life*, Bristol University Press, 2021.

责编 / 陈璐颖

The Technical Myth and Action Choice of Algorithmic Governance

Qiu Zeqi

Abstract: Due to their pervasive and profound influence on the human society and life, algorithms have become an artificial environment comparable to the natural environment. Their advantages and disadvantages as well as the continuous emergence of algorithmic infringement have pushed mankind into the era of algorithmic governance. The influence of algorithms has been increasing, so has their complexity. Algorithms have evolved from the ancient calculation to the self-iteration and evolution of the real-time relationship model, which makes the process governance of algorithms not only technically infeasible, but also costly in management. The exploration and practice in the field of algorithmic governance in China, the United States and Europe show different patterns. The United States has formed an accountability model dominated by the government and the third party with an aim to prevent algorithmic infringement; focusing on data protection, Europe gradually follow the example of the accountability model of the United States; China implements governance from the perspective of Internet security, but has not yet established an algorithmic governance model with a legal logic and clear operation path. Moving towards the accountability model seems to be a possible and effective action choice for China's algorithmic governance.

Keywords: algorithmic governance, algorithmic technology, accountability model