

## 算法社会的“囚徒”风险

彭 兰

**摘 要** 今天算法的应用正在全面普及,一个算法社会正在到来。人们在享受算法带来的便利的同时,在某些方面也面临着成为算法“囚徒”的风险。人的认知、判断与决策可能会受制于算法,人的社会位置也会因算法偏见、歧视以及其他原因受到禁锢。在一些数字劳动平台,算法在隐性控制着劳动者的劳动,算法、大数据及其他新技术也可能增强对人的监控。但反思算法社会的“囚徒”风险,并不意味着对算法的拒绝,加强算法开发者的技术理性和算法伦理培养,提高算法应用者的算法素养,我们才能更好地应对算法社会的挑战。

**关键词** 算法;算法社会;算法伦理;算法素养

**DOI** 10.16602/j.gmj.20210001

近年新闻传播学界对于算法的关注越来越多,但聚焦于算法与信息传播特别是信息茧房的关系的研究居多,然而随着大数据、人工智能技术的发展,算法已经开始全面进入并影响着我们的生活,在某个角度看,一个算法社会正在到来,因此,对于算法的研究,也需要更广阔的视角。

在算法专家看来,算法是“一种有限、确定、有效并适合用计算机程序来实现的解决问题的方法,是计算机科学的基础”(Sedgewick & Wayne, 2012)。通俗地说,算法可以看作用计算机程序实现的、基于数据分析、面向特定目标的一套指令或方案。今天与人们生活紧密相关的算法也是多种多样的,既包括个性化推荐算法、决策算法,也有各种平台的算法、各种管理目标下的治理算法等。

算法的广泛应用是必然的,我们也必须面对算法带来的挑战 and 风险,从个体角度来说,在享受算法带来的各种便利的同时,是否会受到算法的钳制,甚至成为算法的“囚徒”,是需要思考与警觉的。

### 一、人的认知、判断与决策是否会受制于算法?

从本质上来说,算法是一种中介,它基于特定目标下的计算模型,在人与现

彭兰:中国人民大学新闻与社会发展研究中心研究员,中国人民大学新闻学院教授。

实世界之间构建了一种数据化的界面,也因此,它会对人的认知以及在此基础上的判断决策产生影响。

### (一) 推荐算法对人的认知的影响

个性化推荐算法虽然因为近年来一些基于算法的内容平台的兴起而受到关注,但事实上个性化推荐早就进入网络,搜索引擎、电子商务平台等都早已采用推荐算法。

目前主要的个性化推荐算法包括基于内容的推荐、协同过滤推荐、基于标签的推荐、社会化推荐、基于深度学习的推荐、基于知识的推荐、基于网络结构的推荐、混合推荐等。

基于内容的推荐是指根据用户选择的对象,推荐其他类似属性的对象(许海玲等,2009)。在个性化推荐中,这是一种常见的推荐机制。协同过滤推荐(collaborative filtering)是根据用户兴趣的相似性来推荐资源,把和当前用户相似的其他用户的意见提供给当前用户,从而使用户发现新的感兴趣的资源(邢春晓等,2007)。基于标签的方法通过分析用户的标签偏好、物品的标签特征,基于二者的相似性为用户进行物品推荐,其本质是通过引入标签,形成用户—标签—物品三元关系(孔欣欣等,2017)。而近年来开始受到关注的社会化推荐主要是根据用户之间的社会关系信息构建用户之间的社会化关系网络,根据这种社会关系和已知用户兴趣模型,向用户进行推荐(孟祥武等,2015)。基于深度学习的推荐算法则可能更为“聪明”,如通过建模用户的历史序列行为来预测用户的下一个动作、挖掘用户的背景信息以进行更全面的推荐等(刘君良、李晓光,2020)。可以预期,未来的推荐算法还会有更多的新思路。

推荐算法之所以在今天成为一种互联网广泛应用的技术,其核心动力在于解决海量信息(或产品)与用户之间的供需适配问题。对于用户来说,是为他们发现符合自己需要的信息(或产品);对于生产者来说,是为内容(或产品)找到合适的用户。

从内容或产品推荐角度看,作为中介的算法本身就是为用户提供一个过滤器,这种过滤在减少用户的认知负担的同时,也可能局限用户的视野。

美国学者帕里泽很早就关注到了搜索引擎的个性化推荐带来的信息“过滤泡”问题。他指出,个性化过滤器会用两种方式打破我们在强化现有想法和获取新想法之间的认知平衡:其一,它使我们周围充满着我们已经熟悉(并且已经认可)的想法,导致我们对自己的思维框架过于自信;其二,它从我们的环境中移除一些激发我们学习欲望的关键提示(帕里泽,2020,p. 65)。他还认为,个性化推荐限制了我们的解答视界,即寻找问题解决方案的空间大小,也会限制人的创新性(帕里泽,2020,p. 72)。

而国内研究者则更多地从信息茧房的角度探讨算法是否会带来人们的视野狭窄、态度与立场固化等问题。但研究者的判断却并不一致。虽然有很多研究者担忧算法会造成信息茧房,但也有研究者认为信息茧房还不是一种科学概念,目前也还不能证明算法与信息茧房之间存在确切关系。

无论是用过滤泡还是用信息茧房这样的表达,也无论算法与信息茧房之间的关系是否被证实,至少我们可以看到,从算法本身的原理来说,算法的确会带来过滤,这种过滤不可避免会在一定程度上影响人们对外部环境的认知。

在信息爆炸的今天,信息过滤是必然的选择。即使是传统媒体也会对信息进行过滤,在媒体上构建拟态环境。如果媒体对内容的公共性有合理的评估,如果媒体秉持“客观”“平衡”等立场,那么,媒体的拟态环境仍有助于人们全面了解现实社会。专业的媒体在进行信息过滤时,重点考虑的不是人们想要知道什么,而是从社会环境传达与感知角度判断人们应该知道什么。媒体的信息选择,通常也会考虑到内容的平衡与多样性。

但在目前的算法设计中,内容推荐算法主要是参照人们的习惯和相似人群的兴趣,也就是关注人们想要获得什么,在某种意义上也就是在顺应人们的认知心理中惰性的一面,顺应甚至可能强化人的选择性心理。虽然从信息获取效率的角度看,这样的算法可以帮助人们以更小的成本获得与自己偏好、需求更吻合的信息,但是,推荐算法是否只能顺从人们的惯性与意愿,是一个值得思考的问题。信息传播的一个重要目标是社会整合,促成不同人群连接,促进公共对话,这就要打破个人的作茧自缚,因此,内容推荐算法也需要兼顾个性化满足与公共整合双重目标,这也应是算法未来努力的方向。

除了个性化推荐算法对个体的影响,内容推荐算法也会在其平台上营造一种整体的拟态环境。这种拟态环境能否全面、真实地反映现实社会,也与算法的设计思路相关。但今天平台算法过分倚重流量的思路,容易带来马太效应,加重信息环境的不平衡,也使拟态环境与现实环境的偏离加大。

技术界对推荐算法的评测指标主要包括:准确度(用户对推荐内容满意度)、排序合理性、覆盖率、多样性(用户间的多样性和用户内的多样性)和新颖性等(朱郁筱、吕琳媛,2012)。目前提出的评测指标更多的是面向电子商务平台中的商品推荐系统,准确性因此被放在第一位,这种准确性主要也是考虑与用户的惯性、需求的吻合度。但是,对内容推荐算法的评估,还缺乏相应的标准。虽然人文社科领域的研究者呼吁,要在内容推荐算法设计中体现“专业价值观”,但这些价值观具体如何嵌入算法中,仍然是一个很大的挑战。

或许在内容推荐算法的衡量标准中,多样性需要被放到更重要的位置。这种多样性,不仅是内容主题的多样性,还需要包括态度立场的多样性、内容生产者的多样性等。这也是新闻的客观、平衡的专业价值观的具体体现。

内容推荐算法一方面有必要从个性化推荐这一内容消费者视角向内容生产者的视角扩展,也就是说,内容推荐不仅是面向内容消费者的,也应该是面向内容生产者的,有必要通过算法设计让更多内容生产者生产的高质量内容得以传播,特别是使那些具有重要公共价值的内容到达更广的人群。

另一方面,从用户角度看,即使算法在未来能更好地实现内容推送的多样性、个性化与公共化内容的平衡,但如果人们把对信息的选择权完全交给算法,每天都只是等着算法投喂的信息,也会导致他们越来越失去自主性与判断力。

除了推荐算法外,社交机器人对人的认知的影响也是算法影响的另一种表现。社交机器人是指在社交媒体中,由人类操控者设置的、由自动化的算法程序操控的社交媒体账号集群(郑晨予、范红,2020)。在很多社交平台,社交机器人在算法的控制下自动生产着各种内容,这些内容混杂于人生产的内容中,很多时候也不能为一般用户所辨识,因此,通过社交机器人,社交平台的信息环境容易被操控,而这种信息环境也会作用于用户。

在今天的智能化内容创作中,也有算法的作用,无论是文字报道,还是视频,或是其他作品。这些作品在外观上与人创作的作品相似,有时甚至没有差异,但这些作品也只是算法建构的一种认知界面。算法虽然在模拟人的创作思维,并且有可能在某些方面打破人的思维套路,将人带到一些过去未尝涉足的认知领域,但算法本身也有其局限性,它们只能从某些维度反映现实世界,它们对世界的反映仍是相对“平面”的,如果我们总是通过算法构建的界面去认识世界,我们认识世界的方式会越来越单调,也会失去对世界的完整把握能力。

## (二) 算法对人的判断、决策力的抑制

今天很多时候,我们也会依赖算法作出判断与决定,甚至进行大的决策。

对于个体来说,当他接受算法推荐的内容、产品时,某种意义上也是在借助算法作出判断与决策,也就是把对内容与产品的价值判断建立在算法的评价上。

使用导航软件,是在依赖算法进行路线的判断与选择,未来的无人驾驶,更是取决于算法的判断,在这方面,很多时候算法的确可以帮助人们作出正确甚至更优的决策。

在大数据和人工智能技术的支持下,算法对组织机构或一些行业性的决策的作用也越来越显著。例如,银行在对企业或个人发放贷款进行信用和风险评估时,可以参照数据和算法分析的结果。企业在聘用员工时,可能会借助算法进行判断。在医疗领域,智能影像分析系统、疾病诊断系统在帮助医生作出诊疗决策。在法律系统,人工智能也在尝试直接参与决策或进行局部裁判,如进行再犯风险评估、嫌疑人逃脱可能性判断、合理量刑测算等(栗峥,2020)。城市

交通和其他方面的管理也会越来越多地依赖算法。

算法对生产决策的影响也在显现,内容生产行业便是典型。通过各种排行榜来决定内容生产的选题方向或进行内容产品的策划,是今天媒体依赖算法进行内容生产的一种表现,因为各种排行榜也是由一定的算法形成的。除了排行榜,媒体还可以借助更复杂的数据分析与决策算法。例如,2020年B站跨年晚会不仅成功地吸引了B站的用户,也成功地实现“破圈”,这与应用相关数据及算法进行表演嘉宾、表演节目的遴选不无关系。正在兴起的计算广告,其核心思想也是将广告的创意、形式选择、目标人群选择、广告分发与互动等广告生产的全流程建立在数据与算法的基础上。

算法对其他经济活动、经营决策的影响也在日渐深入。例如,在共享经济平台(如网约车平台)上,生产者与消费者直接进行动态、多变、复杂的网状连接和点对点交易,需要依赖平台企业所设计、维护和运营的强大算法(姚前,2018)。企业的产品开发,可以利用算法进行规划、评估,对产品的定价、运营效果的评价以及运营的实时调控等,算法也有其独特的作用。

算法在某些方面的决策中之所以重要,一是因为它可以对与决策相关的对象进行历史、现状甚至未来趋势的数据分析(包括大数据分析);二是因为它可以建立决策模型,基于这种模型对各种可能性进行分析,寻求最优解决方案。算法以过滤信息、建构模型为手段,降低认知负担、提高认知效率(蒋舸,2019),因此,在决策速度与效率甚至某些决策的准确度方面,算法都有可能形成自己的优势。

但这是否意味着人们应该把所有判断与决策都交给算法?

“人们之所以需要借助算法来解决问题,是因为需要借助认知模型将认知负担控制在合目的的范围内。”(蒋舸,2019)既然算法是一种认知模型,它就是对现实世界的一种抽象和简化,很多时候,它只是反映了典型的对象,而不是事物的全部。多样性的世界并非都能通过数据来描绘或计算。算法可能会在某些层面较好地描述与解释现实社会,但在另一些层面则无能为力。完全依赖算法,一些时候就会形成错误的判断与决策。

今天人们的创新,也是一种判断与决策的过程。算法虽然打破了人的某些旧套路,但也会形成一些新套路;如果人们的决策越来越陷入算法营造的套路,那么人类的想象力与创造力也会萎缩。

算法决策主要依赖对事实的判断,但决策过程往往还需要加上情感、道德伦理等其他方面的判断。有法律研究者认为:“司法裁判不是程序化的理性计算,而是事实与价值的复合体(实体上)和技术应用与民主过程(程序上)的统一体,它要借由事实查明和法律适用这一过程体现人性温度、彰显人文关怀。”(陈敏光,2020)这种观点也可以适用于其他一些决策领域。

算法决策中的伦理问题,在今天更是一大困扰,也是未来人工智能发展的重要关切。有研究者指出,人工智能应该包括有伦理影响的智能体、隐含的伦理智能体、明确的伦理智能体、完全的伦理智能体,人工智能在与人的交互中会呈现出某种拟主体性(段伟文,2017);但也有学者认为,人工智能和机器人无法处理开放性情境中的实践伦理问题(蓝江,2018)。未来算法能否更好地解决各种伦理困境还未可知,但即使我们对它在伦理判断中的能力寄予希望,也不能完全把伦理判断交给机器。

另外一个更基本的问题是,我们如何判断算法的可靠性。数据的质量、算法模型的合理性等,都会影响到算法结果的质量。基于数据的算法虽然看似客观,但其实也隐藏着很多主观因素,这些主观因素也会对算法的可靠性产生干扰。

因此,算法社会对人的决策、判断力带来的挑战是双重的:一方面我们要防止把所有决策、判断都交给算法,我们要判断算法在哪些方面可以帮助我们做出更好的决策,而在哪些方面算法可能把我们带入歧途;另一方面,即使在很多时候要参照算法,我们也需要有能力判断算法本身是否有缺陷,算法依据的数据是否可靠,算法是否存在偏见,算法提供的结果是否合理、准确。在没有这种判断力的前提下,盲目依赖算法,也就难免会陷入各种陷阱。

## 二、人的社会位置是否会被算法禁锢?

除了认知与判断可能受制于算法外,人们的社会位置也可能被算法影响。

这首先与算法偏见与歧视相关。算法偏见与歧视既可能源于算法本身的设计,也可能源于算法所依据的数据。

人工智能是人类思维的映像,人类在面对某些问题时采取的“范畴化倾向”的认知态度,也在人工智能的算法流程中体现出来(卜素,2019)。而在这种“范畴化倾向”中,就隐藏着人类思维与文化中的偏见。或许很多算法设计者并没有意识到算法中的偏见或歧视问题,他们只是遵循着社会文化或自己的思维惯例来进行算法设计,这个过程将过去隐性的偏见显现出来。如有研究者指出:“人类文化是存在偏见的,作为与人类社会同构的大数据,也必然包含着根深蒂固的偏见。而大数据算法仅仅是把这种歧视文化归纳出来而已。”(张玉宏等,2017)

在数据层面看,“与算法偏见相关的数据偏见是历史数据本身不可避免的产物。而 AI 等在进行学习训练的时候,恰恰是依据这些有偏见的历史数据,自然其分析的结果带有了很明显的偏见。利用历史数据进行训练的算法程序可能会延续甚至加剧基于种族或性别的歧视”(杨庆峰,2019)。在进行数据挖掘的时

候,常常会把不同的人群进行分类,这种分类也会继承甚至加大不平等体系(林曦、郭苏建,2020)。

虽然在推荐算法中,会存在着算法偏见、歧视等问题,但相对来说,其影响主要是在人们的认知层面,但在决策算法中,偏见以及歧视则有可能影响到人们的权利、社会位置及流动。

目前受到较多关注的算法歧视,如就业歧视、信用歧视、投资歧视等,就在很大程度上关系到人们的社会位置及流动可能。“在算法决策中,个体被赋予一种新的身份,即‘算法身份’,一旦将个体算法身份贴上某种易于被歧视的标签,就会产生双重累积劣势。”(刘培、池忠军,2019)人们往往因为自身的身份和原有的社会位置等,而被决策算法打下某种标记,原来处于优势地位的人往往会获得有利的标记,因而有获得更多资源和向上流动的可能性,而原来处于不利地位的人,则因标记而处于更不利地位,从而失去就业、获得投资等相应机会。算法可以对人们的历史记录进行追踪,将各种平台的数据相关联,将人们的背景翻个“底朝天”,这使得人们更加难以逃离既有社会位置的束缚。“当算法将虚拟世界和现实世界的个体相联结,将某一个体的过去、现在和未来相联结时,一次性的不公可能暗藏着对个体结构性的歧视锁定。”(张欣,2019)

如果没有足够的规制,未来算法歧视可能还会在教育、医疗等其他领域体现。

在一定意义上,算法的偏见和歧视不能完全避免。对于算法的设计者来说,需要一定的机制包括法律上的约束,尽可能减少算法偏见、歧视的产生;而对于一般人来说,则需要意识到算法偏见、歧视在哪些方面存在,它们是如何对个人产生影响的。

除了算法偏见与歧视外,算法对人们的社会位置的禁锢还会通过其他方式实现。前文所说的推荐算法带来的认识局限,也会在一定程度上影响人们对其他群体的了解。算法在连接相似人群的同时,也可能会造成不同人群之间的区隔。在这个意义上,算法也容易将人们限定在某个特定“圈”“层”中。

算法社会带来的另一种社会分化是信息技术的贫者与富者的分化。“算法社会一定是科技精英社会。少数人会成为主宰,而大多数人只能顺从。”(於兴中,2018)在某种意义上,算法会加剧数字鸿沟,信息技术的贫者不仅与算法权力无缘,也会在他人的算法权力的控制下,被困在自己的社会阶层中。

### 三、人的劳动是否会被算法隐性控制?

2020年9月,一篇题为《外卖骑手,困在系统里》的特稿,引起了很多人的关注。报道中提到,外卖平台有一种算法,从顾客下单的那一秒起,系统便开始根

据骑手的顺路性、位置、方向决定派哪一位骑手接单,订单通常以3联单或5联单的形式派出,一个订单有取餐和送餐两个任务点,如果一位骑手背负5个订单、10个任务点,系统会在11万条路线规划可能中完成“万单对万人的秒级求解”,规划出最优配送方案(赖祐萱,2020)。虽然对于顾客来说,这种算法可以使他们在最短的时间里拿到外卖;但对于骑手来说,这却可能意味着不断提高的时间压力和劳动强度。

有研究者指出,送餐时间的不断缩短与算法对送餐员的规训密不可分。数字平台通过算法中介了劳动和消费的关系,通过建构高效、及时等时间话语来赢得资本市场,但同时也对外卖送餐员实行了算法管理下的时间规训和时间操控。送餐平台依靠算法精细化的管理,将传统情境下的情感劳动付出合理化和规范化,进一步实现了对于送餐员的纪律规训(孙萍,2019)。

其他研究者也有类似的判断。有研究发现,相比传统劳动中的雇主控制、等级控制、技术控制和科层制控制,外卖骑手的劳动是由平台控制的。平台系统将劳动过程精确到可计算的程度,实现了对劳动的高度控制和精准预测,这很大程度上源于平台系统背后的数据、算法和模型的支持(陈龙,2020)。甚至有些平台通过算法升级,使原本由骑手自己决定工作内容与工作数量的状况转变为依靠算法强制分配。同时,拒绝算法的强制分配会付出隐秘但更严重的代价——可接收的订单总数量大幅下降(叶韦明、欧阳荣鑫,2020)。

数字平台上的内容生产,是一种更直接的数字劳动,这种数字劳动也在很大程度上受到平台规则的影响。平台规则会调节供需关系和刺激劳动,而平台规则往往也需要付诸算法。

有研究者对网络文学的生产机制研究指出,“付费阅读模式”已成为文学网站通行的营利规则,为了获取读者的关注度,也为了得到更多的金钱回报,网络作家开始越写越长,写作不再是自由的、个性化的自我呈现,迎合读者和市场需求、维持读者的阅读快感和引发阅读欲望成为网络文学作品生产的内在逻辑。“文艺青年”逐渐演变为没有固定劳动合同、按照单件作品的文字量获取报酬的“计件工”(蒋淑媛、黄彬,2020)。

同样,在短视频、视频直播等平台,算法在激励用户参与内容生产的同时,也在无形中控制着他们的劳动付出和回报,甚至异化着他们的劳动目标。

平台算法之所以能对人们的劳动产生直接的控制,一个重要原因是平台将劳动者与消费者直接连接,消费者可以通过平台反馈、评价劳动者的劳动;劳动者的结果可以直接量化,并成为评价劳动的主要指标,而这是算法最适合做的事。通过量化结果来控制劳动,看上去劳动者的劳动过程变得自由了,但为了获得更好的劳动结果,他们会付出更多的劳动,包括情感劳动,通过各种方式取悦消费者,以获得额外的肯定。



对于内容生产者来说,流量成为最基本的评价指标。有研究者指出,数字空间对流量的追逐带来了交往资本主义,讯息的另一个价值变得更为重要,这就是贡献流量。重要的不是内容,而是流量。在此流量和彼流量之间没有实质内容上的差别,只有流量计算大小的差别(蓝江,2019)。将对内容本身的信息量、质量优劣的评价简单化为对流量的评价,这是新媒体时代传播发生的重要变化,平台对流量进行的数据统计与在传播界面中的直接呈现,以及各类基于这些数据的排行榜、指标体系,造就了传播评价体系的算法化。这不仅给劳动者施加了压力,也有可能使他们的内容生产的出发点发生偏移,他们不再是作为内容生产者作出独立判断,而是要时时顾及内容消费者的心理与需求,在社交媒体中,他们尤其需要考虑作为“人媒”的社交网络对内容的需求。虽然在一定程度上,面向消费者的考量可以使内容生产者的生产更有目的性,与市场需求产生更多呼应;但另一方面,过于单一的评价算法体系也会使得简单的“鼠标投票”机制取代专业的评判机制,内容生产者自身的专业判断力因此可能被削弱。

当影响力或舆论声量等都靠流量来衡量时,一种另类的数字劳动者——网络水军也就应运而生。利用碎片化时间进行有偿数字劳动的网络水军,其主要工作就是人为制造各种数据(如刷分、刷评论、刷单),以影响相关评价;在某种意义上,网络水军是量化制度设计的“社会症候”(吴鼎铭,2019)。水军这样一种数字劳动者本身的劳动报偿与算法紧密相关,他们又改变了他人的评价体系中的相关数据,对他人的劳动产生影响。

平台还有可能将劳动改头换面,将一些劳动变成舒尔茨所说的“玩劳动”(play-labor、playbor),实现生产性和休闲性的统一(李弦,2020),使人们更心甘情愿地参与到数字劳动中。甚至在克里斯蒂安·福克斯看来,今天所有的数字平台的用户实际上就是运输工人,而且是免费替大平台和广告商运输商品意识形态的工人,都是一种数字劳动者(蓝江,2019)。虽然新媒体用户的“劳动化”是多重力量推动的,但算法在其中的作用也不容忽视。

平台的各种机制通过数据与算法模型体现出来,平台的控制最终会演变为劳动者的自我约束与激励,在这种自我约束与激励下,一些劳动者会变成“永动机”。

而在很多劳动领域,即使没有平台的存在,面对各种量化考核指标,人们也不得不给自己加压,但人们的努力却会带来水涨船高,考核指标不断升级,人们的压力不但没有减轻,反而进一步加强,这也是2020年备受关注的内卷现象的一种典型表现。虽然内卷这个词近年才流行,但内卷现象早就存在。用各种简单、表面的考核指标衡量劳动者(包括脑力劳动者)的成果,与流量的统计类似,也可以说是一种简单的评价算法,而在算法时代,这样的评价思维似乎还有愈

演愈烈之势。

当然,也有一些平台试图通过算法的调节,缩小劳动者流量的“贫富差距”,使更多的劳动者的劳动得到关注,获得更多回报,虽然这并不能从根本上解放劳动者,但也说明,算法本身存在多种可能性。算法本身是中性的,算法既可能禁锢劳动者,但应用得当,也有助于为劳动者松绑、减负,只是对此今天我们还没有充分的思考和行动。

#### 四、算法社会对人的监控会强化吗?

随着人工智能应用的发展,人工智能为国家权力对个人和社会的监管提供了数据基础,“治理体系的算法化”开始萌生(王小芳、王磊,2019),一种集成运用互联网、物联网和人工智能技术构建的以“智慧管理器”为中介系统的智慧社会也正在到来(何明升,2020)。算法在这样的智能管理中也扮演着核心的角色。

从积极的方面看,如学者们所指出的,在社会管理中,人工智能在硬性价值引导(经过全面的技术监控让人信守规矩)、社会制度设置及其守持、生活世界的秩序制定与维护上,已经发挥出重要的作用(任剑涛,2020)。智能化社会治理可以通过大数据、人工智能等技术将复杂的社会运行体系映射在多维、动态的数据体系之中,不断积累社会运行的数据特征以应对各类社会风险、提升社会治理有效性(孟天广、赵娟,2018),社会治理规则的算法化有助于提高社会治理的主动性与预判性,让社会治理主体更加主动地预测预警预防社会风险(周汉华、刘灿华,2020)。算法的不透明也不一定是坏事,如果运用合理,不透明的算法也有可能弥补社会裂痕,维持社会共识(丁晓东,2017)。

但在包括算法在内的人工智能技术进入社会治理的过程中,个体受到各种监视、控制的风险也在加大。

算法管理的一个基础是人的数据化。今天对人的数据的收集,不仅涉及人们主动在新媒体平台生产或提供的数据,也涉及很多人们被动提供的数据。数据采集的工具也向各类传感器、可穿戴设备拓展,它们对于人的数据的采集进入深层,人的现实行为数据、生理数据等成为收集对象,其中很多是人们不愿意透露的甚至是涉及个人隐私的数据。

对人的数据的广泛采集虽然表面上看给生活带来了一些新的便利,例如,人脸识别加快了人们支付、安全审核等的速度,但这些便利后面往往隐藏着巨大的风险;而个体可能对这些风险毫无知觉,即使他们能意识到风险,很多时候也无法与数据的收集机构相抗衡。强制人们进行各种形式的数据化,用个人数据兑换各种服务便利或权利,已经成为算法社会的一个普遍事实。

在数据收集的基础上,算法可以进一步对个体进行计算,从而发现数据背后个体的更深层的秘密,并以此对其进行控制。算法对人的控制是全程的控制,人们的每一个活动和行为都可能成为当下算法的依据,也会累积起来影响到未来的算法计算结果。“表面上看起来不起眼的一次行为活动,实际上已经被后台看不见的算法程序演算过无数遍,任何选择都是符合算法的选择,最终,我们看似自主的行为全部在算法治理的彀中。”(蓝江,2020)

算法隐含着各种社会规则,与算法相关的评分机制,则把人们对规则执行的结果量化出来。因此,算法与评分很多时候结伴同行,“这种评分机制可以帮助汇集社会主体的日常活动,将弥散化的社会意识与价值判断进行一定程度的集中,形成公意并强制执行”(胡凌,2019)。在一定意义上评分制强化了人们对社会规则的认识与遵守,激发了人的自我约束,“基于数据—算法—后果模式下的信用评级裁量决策系统取代法律—行为—后果模式下的法律构成要件分析框架。社会信用的生成源自数据提取,是单向的自我问责过程,而非基于双向法律关系”(虞青松,2020)。基于评分的奖惩简单、直接,有时也是有效的。

评分机制在某些时候对于社会风险控制是有价值的,但不可否认的是,它可能形成对个人隐私的侵犯以及对算法控制权的滥用。如研究者指出,评分权力可能会造成道德与社会规范发挥作用空间的压缩,权力的实施方式更加深入和动态化(胡凌,2019)。社会信用监管者能够进入以往传统权力所无法企及的私人领地,将行为人的私人空间(包括心理状态、行为道德等)统统纳入社会信用系统的监管之下(虞青松,2020)。2020年某地方政府提出的“文明码”等之所以受到质疑,也在于其试图滥用管理权力,通过算法入侵道德评价这样的私人空间。

算法下的评分制不仅为管理机构的监控提供了支持,也为人们之间的相互评价、监督提供了基础。算法技术的发展也可以使人人成为他人数字人格的观察者、执法者、裁判者(虞青松,2020)。虽然在某些情况下,这种评分可以为人们在网络空间的互动(特别是交易)提供安全评价依据,便于人们进行风险判断,但与此同时,用户间的相互评分的权力也可能被滥用,用户间的相互监控会使个体面临更大的压力。

不仅算法治理需要个人数据,各类企业也在收集个人数据,以获得市场的分析与运营基础。对于企业来说,算法是用来实现利润最大化的工具;而对于算法而言,个人信息与市场数据就是能量棒(扎拉奇、斯图克,2018)。与此同时,“算法与商业资本结合形成监视资本主义,用户被嵌入数据生产链条,变成被算法支配调控的客体”(张凌寒,2019)。企业对用户的控制不仅表现为对用户的个人信息与数据的利用与控制,还表现为对他们需求与行为的控制,算法

在不断挖掘用户的潜在需求,甚至诱导出他们的需求,助推消费主义倾向。

在外部力量通过算法等强化对个体的控制的同时,个体自身在数据和算法的导向下,也可能会在自我传播或社会互动中强化自我审查。算法的监控,也会内化为人的自我规训。

学者段伟文指出,作为一种新的智能社会形态的数据解析社会已悄然而至,其运作和治理基础是用数据和智能算法来分析人的行为,但这种智能监测属于揣测性的认知,有可能误读和不当干预被认知对象的能动性,急需为可能出现的技术滥用设定边界(段伟文,2020)。也有研究者呼吁智慧社会背景下的以数字人权为代表的第四代人权(马长山,2019)。但这些目标的实现,还有很长的路。

## 五、结语：我们如何对抗算法的囚禁？

尼尔·波兹曼在《娱乐至死》一书前言中提到了《1984》和《美丽新世界》暗喻的两种警告。“奥威尔警告人们将会受到外来压迫的奴役,而赫胥尼则认为,人们失去自由、成功和历史并不是老大哥之过,在他看来,人们会渐渐爱上压迫,崇拜那些使他们丧失思考能力的工业技术。”(波兹曼,2004)算法在某种意义上是这两种风险的结合,一方面,算法对人的算计越准,就意味着它对人的了解越深,因此,对人的监视与控制也可能越深;另一方面,当算法对人的理解越深,对人的服务越“到位”,人们从中获得的满足也越多,而对算法的依赖、依从也会越多。当算法渗透到社会生活的各方面,人对它的依赖成为惯性,人对算法带来的囚禁也可能会越来越浑然不觉。

“算法社会将自由与枷锁的张力推向了极致”(齐延平,2019),算法一方面在促成人的某些能力的解放与扩张,另一方面又用某些方式实现着对人们的禁锢。但是,当我们深入反思算法对人的各种囚禁时,我们的目的并不是将算法拒之门外,这一点正如我们对待汽车的态度。

汽车进入我们生活,带来了正向与负向的双重影响,但人类的解决方案不是禁止汽车的使用,而是通过对驾驶技能的培训,以及严格的交通法规制定与实施等,来尽可能减少其可能产生的危害。同样,当算法成为一种广泛应用的技术,在很多方面可能带来对人的囚禁风险时,我们也不能简单禁止算法的使用。除了在法律、制度等层面作出必要的调整外,也需要面对算法社会的新特点,培养不同主体的相应素养与能力。

对于算法的开发者来说,新的技术理性、算法伦理的倡导与培养,尤为关键。

近年来对算法等智能化技术的批评中不乏技术理性批判的视角。尽管反

思与批评是必要的,但也如有学者所指出的有些技术理性批判存在一个误区,那就是把技术理性等同于工具理性,认为技术理性张扬必然导致价值理性衰微(赵建军,2006)。也有些研究虽然没有涉及技术理性这样的概念,但也习惯性地将技术思维与工具理性画了等号,也就是有意无意地默认技术一定是以工具理性为导向的。然而,正如一些研究者意识到的,技术理性本应是工具理性与价值理性的内在统一,只是由于这两种理性之间的内在张力才使技术理性始终处于内在的矛盾运动中。随着人类技术实践活动的扩展和深入,技术理性内在的矛盾才以一种单向度的、异化的形式呈现出来,即工具理性压倒价值理性、技术的价值理性萎缩成极度膨胀的工具理性的单纯附属物(刘祥乐,2017)。

在今天,有必要重新理解技术理性应有的“题中之意”,在算法开发者群体中倡导价值理性与工具理性、技术思维与人文精神的融合,而不是将算法推向工具理性的极端。在此基础上,充分探讨算法伦理的目标、原则与实现路径,并使之成为算法开发者的制衡。

对于算法的使用者来说,算法时代带来了对人的素养的新要求。除了一般的媒介素养、数字素养外,也需要一定的算法素养。

就像媒介素养一样,倡导算法素养的前提,不是简单地将算法认定为坏的东西,让人们一概排斥算法,而是要让人们意识到在今天这个时代,算法无法避免,因此,重要的是要理解不同类型的算法是如何运作的,算法在哪些层面影响着我们的认知、行为、社会关系,影响着我们的生存与发展,在此基础上学会与算法共存,对抗算法的风险,更好地维护人自身的合法利益与地位。

面对一个无可回避的算法社会,我们只有提高对算法的认识与驾驭能力,才能成为算法的主宰者,而不是其“囚徒”。

## 参考文献

- 阿里尔·扎拉奇、莫里斯·E. 斯图克(2018):《算法的陷阱:超级平台、算法垄断与场景欺骗》(余潇译),北京:中信出版社。
- 卜素(2019):人工智能中的“算法歧视”问题及其审查标准,《山西大学学报(哲学社会科学版)》,第42卷第4期,124-129页。
- 陈龙(2020):“数字控制”下的劳动秩序——外卖骑手的劳动控制研究,《社会学研究》,第6期,113-135页。
- 陈敏光(2020):司法人工智能的理论极限研究,《社会科学战线》,第11期,194-203页。
- 丁晓东(2017):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》,第29卷第6期,1609-1623页。

- 段伟文(2017):控制的危机与人工智能的未来情境,《探索与争鸣》,第10期,7-10页。
- 段伟文(2020):面向智能解析社会的伦理校准,《上海交通大学学报(哲学社会科学版)》,第4期,27-33页。
- 何明升(2020):智慧社会:概念、样貌及理论难点,《学术研究》,第11期,41-48页。
- 胡凌(2019):数字社会权力的来源:评分、算法与规范的再生产,《交大法学》,第1期,21-34页。
- 蒋舸(2019):作为算法的法律,《清华法学》,第13卷第1期,64-75页。
- 蒋淑媛、黄彬(2020):当“文艺青年”成为“数字劳工”:对网络作家异化劳动的反思,《中国青年研究》,第12期,23-30页。
- 孔欣欣、苏本昌、王宏志、高宏、李建中(2017):基于标签权重评分的推荐模型及算法研究,《计算机学报》,第40卷第6期,1440-1452页。
- 赖祐萱(2020年9月8日):外卖骑手,困在系统里,获取自<https://mp.weixin.qq.com/s/Mes1RqIOdp48CMw4pXTwXw>。
- 蓝江(2018):人工智能与伦理挑战,《社会科学战线》,第1期,41-46页。
- 蓝江(2019):交往资本主义、数字资本主义、加速主义——数字时代对资本主义的新思考,《贵州师范大学学报(社会科学版)》,第4期,10-19页。
- 蓝江(2020):生命档案化、算法治理和流众——数字时代的生命政治,《探索与争鸣》,第9期,105-114页。
- 李弦(2020):数字劳动的研究前沿——基于国内外学界的研究述评,《经济学家》,第9期,117-128页。
- 栗峥(2020):人工智能与事实认定,《法学研究》,第1期,117-133页。
- 林曦、郭苏建(2020):算法不正义与大数据伦理,《社会科学》,第8期,3-22页。
- 刘君良、李晓光(2020):个性化推荐系统技术进展,《计算机科学》,第47卷第7期,47-55页。
- 刘培、池忠军(2019):算法歧视的伦理反思,《自然辩证法通讯》,第41卷第10期,16-23页。
- 刘祥乐(2017):技术理性批判的逻辑嬗变及其困境:从卢卡奇、马尔库塞到哈贝马斯,《内蒙古大学学报(哲学社会科学版)》,第49卷第5期,43-49页。
- 马长山(2019):智慧社会背景下的“第四代人权”及其保障,《中国法学》,第5期,5-24页。
- 孟天广、赵娟(2018):大数据驱动的智能化社会治理:理论建构与治理体系,《电子政务》,第8期,1-11页。
- 孟祥武、刘树栋、张玉洁、胡勋(2015):社会化推荐系统研究,《软件学报》,第26卷第6期,1356-1372页。
- 尼尔·波兹曼(2004):《娱乐至死》(章艳译),南宁:广西师范大学出版社。

- 齐延平(2019):算法社会的治理逻辑,《华东政法大学学报》,第6期,5页。
- 任剑涛(2020):人工智能与社会控制,《人文杂志》,第1期,33-44页。
- Sedgewick, R. & Wayne, K. (2012):《算法》(谢路云译),北京:人民邮电出版社。
- 孙萍(2019):“算法逻辑”下的数字劳动:一项对平台经济下外卖送餐员的研究,《思想战线》,第45卷第6期,50-57页。
- 王小芳、王磊(2019):“技术利维坦”:人工智能嵌入社会治理的潜在风险与政府应对,《电子政务》,第5期,86-93页。
- 吴鼎铭(2019):量化社会与数字劳动:网络水军的政治经济学分析,《现代传播》,第5期,47-51页。
- 邢春晓、高凤荣、战思南、周立柱(2007):适应用户兴趣变化的协同过滤推荐算法,《计算机研究与发展》,第44卷第2期,296-301页。
- 许海玲、吴潇、李晓东、阎保平(2009):互联网推荐系统比较研究,《软件学报》,第20卷第2期,350-362页。
- 杨庆峰(2019):数据偏见是否可以消除,《自然辩证法研究》,第35卷第8期,109-113页。
- 姚前(2018):算法经济:资源配置的新机制,《清华金融评论》,第10期,91-98页。
- 叶韦明、欧阳荣鑫(2020):重塑时空:算法中介的网约劳动研究,《浙江学刊》,第2期,167-176页。
- 伊莱·帕里泽(2020):《过滤泡:互联网对我们的隐秘操纵》(方师师、杨媛译),北京:中国人民大学出版社。
- 虞青松(2020):算法行政:社会信用体系治理范式及其法治化,《法学论坛》,第35卷第2期,36-49页。
- 於兴中(2018):算法社会与人的秉性,《中国法律评论》,第2期,57-65页。
- 张凌寒(2019):算法权力的兴起、异化及法律规制,《法商研究》,第36卷第4期,63-75页。
- 张欣(2019):从算法危机到算法信任:算法治理的多元方案和本土化路径,《华东政法大学学报》,第22卷第6期,17-30页。
- 张玉宏、秦志光、肖乐(2017):大数据算法的歧视本质,《自然辩证法研究》,第33卷第5期,81-86页。
- 赵建军(2006):超越“技术理性批判”,《哲学研究》,第5期,107-113页。
- 郑晨予、范红(2020):从社会传染到社会扩散:社交机器人的社会扩散传播机制研究,《新闻界》,第3期,51-62页。
- 周汉华、刘灿华(2020):社会治理智能化的法治路径,《法学杂志》,第9期,1-12页。
- 朱郁筱、吕琳媛(2012):推荐系统评价指标综述,《电子科技大学学报》,第41卷第2期,163-175页。

## The Risk of Imprisonment in the Society of Algorithm

Lan Peng

*(School of Journalism, Renmin University of China)*

**Abstract** With the application of computer algorithm in many areas, an algorithmic society is coming. When enjoying the convenience brought by algorithm, people also face the risk of becoming prisoner of algorithm. People's cognition, judgment and decision may be controlled by algorithm. People's social position may be restrained because of bias and discrimination of algorithm or other reasons. On the digital laboring platform, algorithm controls people's labor in hidden ways. In the algorithmic society, monitoring on people may increase with the use of algorithm, big data and other new technologies. Rethinking of the risk of algorithm does not mean rejection of it. Enhancement of technological rationality and ethics of algorithm developers, as well as training of algorithmic literacy of users, may help us to meet the challenges in the algorithmic society.

**Key Words** Algorithm; Algorithmic Society; Algorithmic Ethics; Algorithmic Literacy

(编辑:卢 嘉)