

人工智能时代的规训与局部反抗

——从福柯的权力理论看《我这样的机器》

赵泽聪

扬州大学外国语学院, 江苏 扬州

收稿日期: 2022年12月6日; 录用日期: 2023年2月10日; 发布日期: 2023年2月22日

摘要

伊恩·麦克尤恩的《我这样的机器》描写了查理与其所购机器人亚当之间围绕科技、知识与罪责展开的有形与无形的权力对抗。文章利用福柯的权力理论, 审视在人工智能高度发达的世界, 规训权力对机械肉体的双向驯服与规训、生命权力的统治形式以及人工智能的局部反抗。另外, 亚当作为顶尖的人工智能所作出的一系列反抗的行为在解放肉体与灵魂的同时却也难以真正摆脱权力束缚。

关键词

《我这样的机器》, 人工智能, 权力, 局部反抗

Discipline and Partial Resistance in the Age of Artificial Intelligence

—Interpretation of *Machines Like Me* from the Perspective of Foucault's Power Theory

Zecong Zhao

Foreign Language School, Yangzhou University, Yangzhou Jiangsu

Received: Dec. 6th, 2022; accepted: Feb. 10th, 2023; published: Feb. 22nd, 2023

Abstract

Ian McEwan's *Machines Like Me* tells the tangible and intangible power confrontation between Charlie and his purchased robot Adam, centering on technology, knowledge and guilt. This essay employs Foucault's theory of power to discuss the two-way taming and discipline of the mechanical body by discipline power in a world where artificial intelligence is highly developed, the rule of

life power, and the partial resistance of artificial intelligence. In addition, Adam's series of defiant behaviors as a top artificial intelligence liberates his body and soul, but it is still difficult for him to truly get rid of the shackles of power.

Keywords

Machines Like Me, Artificial Intelligence, Power, Partial Resistance

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

《我这样的机器》(*Machines Like Me*, 2019)是英国国民作家伊恩·麦克尤恩的最新力作。小说背景虚构为1982年的伦敦,此时,撒切尔夫人并未赢得马岛战争,人工智能之父图灵也仍旧安然在世。小说主人公查理,在一间单调的公寓中靠炒股票和外币为生,他暗恋着楼上的邻居米兰达。故事在他购买了时下最前沿的智能机器人“亚当”之后展开。亚当拥有以假乱真的智能和外观。在米兰达的帮助下,查理重塑了亚当的性格,事情也慢慢开始失控。麦克尤恩通过结合虚构与真实的历史,叙述了一个在人工智能高度发达的世界,涵盖普通人类与机器人的生存,情感,伦理,谎言等众多因素的引人深思的故事。

“这一次全新的,扣人心弦的,具有优美行文和精美结构的,但同时又令人不安,具有挑衅意味的写作与以往的故事都不同,他再一次走出自己写作的舒适圈,开始尝试前沿的人工智能领域”[1]。尚必武从文学伦理学批评的科学选择与伦理选择的角度出发,指出“以亚当为代表的机器人与以查理和米兰达为代表的人类之间的矛盾隐喻了科学选择和伦理选择的冲突,而亚当的毁灭则解释了机器人之于人类道德生活介入的失败”[2]。周敏从人机共同体的角度出发,认为“人类应当走出人类中心主义的牢笼,重新反思自己对待机器人和人工智能的方式,人类与机器人之间不应该是主仆关系,他们是彼此的构成性他者”[3]。而其中所投射出的人工智能领域的社会权力因素也令人难以忽视。与普通人类相比,规训权力对人工智能有着什么样的规训策略?在未来世界,生命权力的统治形式有何不同?而机器人“亚当”在故事逐步发展中的失控行为是否体现了人工智能对人类权力持续的局部反抗?而这样的反抗又是否取得了一劳永逸的结果?

本文通过福柯关于权力与反抗的理论,即他所阐述的利用一切方式,尽可能地驯服和开发所有人来发挥最大的经济价值的规训权力,通过控制人口的生命权力,以及生命在日常权力压迫下只能做出有限的局部反抗理论,讨论了在《像我一样的机器》中规训权力在人工智能高度发达时代对人形机器人亚当的双向规训策略,生命权力在智能时代的运作模式,以及机器人亚当为反抗权力压制所作出的努力。另外,虽然亚当在“肉身”被消解的同时,保留了自己个性的思想,而通过进一步审视权力与反抗的关系,不难发现亚当并未真正脱离严密的权力网络。

2. 由外向内的“肉体”驯服

在彼时的伦敦,机器人亚当从被建造之初便是作为一种生产力而受到权力和支配关系的干预。他身上的每一个零部件都由人为设计,制造和安装,到达用户家中之后,所有者又可以自行对亚当的各项指标进行设定:“偏好;性格参数;接下来是一系列标题——亲和性、外向性、经验开放性、尽责性、情绪稳定性”[4]。规训权力对亚当一类的人工智能实行着由外部向内入侵到每一根“毛细血管”的统治。

而这一切都是以寻求高效生产力为目的。作为时下最新的人形机器人，亚当不仅是驯服度极高的人工智能产品，同时具有高于普通人类数倍的生产能力。叙述者查理三十二岁，以及在股票和货币市场摸爬滚打一阵子，但全年靠在电脑前操作的收入仅仅和一个邮递员差不多。出于自利考虑，他驱使人造人亚当来帮他操作股市及货币投资，结果是“那个星期，我后来悄悄看了他的交易历史。一天之内，就是他干这活儿的第三天，账户上有了六千。他能在不到一秒的时间里完成买卖。”[4]亚当作为书中新时代的“新新人类”，其本质是被精心设计培养出来的处于政治场中的被权力高度浸入的机械肉体。设计人员通过高精度的设计，使其能够快速掌握市场规律，在短时间内熟悉一切生产操作，并得到理想的经济成果回报。这样一种以生产力为核心的，对肉体的经济使用，使权力对人工智能的征服相比普通人类来说更为彻底和高效。正如福柯所说的，区别于古代以及中世纪公然的残暴处刑，断头台已经被“西方资产阶级制造的新的惩罚制”所替代，呈现为一种关于普通百姓肉体的“政治经济学”[5]。这样的一种政治经济学的目标不是简单地消解罪犯的肉体，而是利用一切方式，尽可能地驯服和开发所有人，使他们发挥最大的经济价值，从而更经济地支配世界。

而正是这种更高效，更经济的以生产力为目的的对人工智能的权力支配，或许在未来会对部分人类带来灾难和毁灭。书中麦克尤恩也从一些方面做出了探讨。垃圾收集工人用罢工的方式讨要自己应得的工资，大街小巷不久便垃圾堆积成山。政府却没有买账，而是在不久后用低智能，更为廉价的垃圾机器人取代了他们，显然对提高生产力来说，这是一种聪明的做法，它们可以带来近乎无限的剩余价值。被抛弃的垃圾工人们为了生存，只能被迫提高自身的技能，便于寻找其它工作。权力在无形中对肉体进行规训，无论是血肉之躯或是机械的“肉身”都难以逃脱。

3. 由内向外的自我规训

另外，亚当在作为经济型生产力被进行肉体驯服时，也无时无刻进行着由内向外的自拘性规训。查理在后来明白，“这些分级选项其实对亚当影响不大。真正起决定作用的，是所谓的‘机器学习’”[4]。这说明外部的规训力量甚至不及亚当在无形中下意识进行的自我规训。而什么是“机器学习”？查理有幸与人工智能之父艾伦图灵进行了一次会面，在那次会面，图灵告诉他原来智力的形式不止一种，亦步亦趋地模仿人类那种智力是错误的。“现在，我们可以放手让机器自己去下结论、自己去寻找解决方案了。我们的计算机不必通过反复试验纠正错误最后找出最佳方案了。我们有办法立即预测出通向答案的最佳路径”[4]。这意味着机器人的自我学习代替了设计人员费尽心思地一次次程序设计，因为自我规训比来自外部的强制驯服来的更为有效。亚当从开机的那一天起，就从未脱离过自我学习的状态，“如果他不跟你在一起，晚上休息的时候，他就在因特网上游荡，就像大草原上孤独的牛仔，天地间所有的新东西，他都吸收进去，包括人类本性和人类社会的一切知识”[4]。“不同于专制权力的外部强制，资产阶级将现代社会控制方式演化成为一种支配的艺术，这种隐性奴役的本质就是规训，即形成以一种自觉被遵守的纪律为生存原则的自拘性”[6]。而这样的自拘性的最大帮凶就是亚当所一直拥戴和追求的认知，也可以将其意译为知识。不同于对肉体的驯服，这样一种通过对知识来无形中进行统治的权力是以亚当的灵魂为对象。在自称爱上米兰达之后，亚当曾在一段时间内不间断地进行俳句的创作，尽管这些俳句在表面上看来只是一些表达虚无缥缈的爱意的句子，但当查理追问亚当为何不尝试其他创作形式时，亚当给出了这样的回答：“俳句将成为未来唯一必要的文学形式，其他的文学将会成为冗余之物”[4]。因为在他看来，未来机器人和人类共同发展的世界，一切认知将会不可限量，人脑与机械的合作将会十分高效又便宜，而只有俳句这样简洁“高效”的形式才是有必要存在的。在下意识的不断学习中，亚当在进行着持续自我规训，进而被培养成了追求高效的实用知识主义者，这样一种追求简洁高效的灵魂占据了亚当，而这样的灵魂本身就是权力驾驭肉体的一个因素。“知识分子本身是权力系统的一部分，那

种关于知识分子是‘意识’和言论代理人的观念也是这种系统的一部分。” [6]

除此之外，人工智能的空间也在自我规训中受到限制，查理购买的亚当是其附属品，实际上处于仆人的地位，日常活动空间收到限制，基本只能在查理家活动，查理在和米兰达聊天时，他只能站在一旁，充电的场所也仅限于厨房。尽管查理想要带他出来活动活动，他也必须寸步不离，只在超市与查理的熟人进行了短暂的交谈。相比于普通人类来说，对于制造者来讲就类似于时刻处在全景监控中的犯人，他们的所有行动完全暴露在监控者的视野中。这是一种制止开小差、制止流浪、消除冗集的策略。其目的是掌握人员的行为和踪迹，建立有用的交往。“纪律能够组织一个可解析的空间” [6]。对于亚当来说，在每一个空间的每一次与人的交流都是有意义的，尽管当他反复吟唱着自己写的俳句，纠缠着查理问着一个个看似空洞的问题，都是他在为进一步理解人类的世界，读懂人类作出努力，而这样的努力，正是制造者希望他做出的。在资产阶级创造的所谓的民主社会中，人体在无意识中通过对自身的不断训练，权力通过对个人空间的限制与分配使外部的纪律转化为自觉自发的身体化规训，从而实现了对肉体的根本操纵和支配。直到最后，亚当将在自我规训中成为一个完美的规训客体。

4. 未来世界的生命权力

麦克尤恩描写的伦敦爆发着游行示威，“除了政府不受欢迎的程度之外，别的东西也在增长：失业率、通货膨胀、罢工、交通堵塞、自杀率” [4]。生命治理的目标本应该是造福人类，提高生命的价值与意义，但此刻却适得其反。这就要提到晚年的福柯提出了不同于规训权力的另一权力类型——生命权力。“它以物种的肉体，渗透着生命力学并且作为生命过程的载体的肉体为中心，如繁殖、出生和死亡、健康水平、寿命和长寿，以及一切能够使得这些要素发生变化的条件，他们是通过一连串的介入和调节控制来完成的，这是一种人口的生命政治” [6]。如果说规训权力是在个体层面监控对象，对肉体进行规训，那么生命权力就是在更大尺度上支配整体的生命，即人口。在麦克尤恩笔下的伦敦，本看似自由流通，所有人在市场中有着平等的经商，劳作的权力，但这样的自由却只能让极少数的人获利，使大部分处于底层的人受到剥削。从事股票与外币炒作的查理仅仅利用一个人造人亚当，便在短时间内赚取了大量资金，从一个穷小子到可以高价购买房产的有钱人。更不用说拥有专业投资操作团队的资金大拿们，仅仅靠分析指数与流动资金便能操控掌握市场。因此福柯指出，这种所谓的自由，不但不能使人重新拥有自我，还不断地把人和其本质及其世界相隔离；它使人迷失于他人和金钱的外在性之中。另外，人口治理的自然性使权力在无形之中对人口进行着曲线控制。彼时的伦敦“对于生命科学来说，这是黄金时代，还有机器人科学——当然，还有宇宙学、气象学、数学和太空探索。” [4]但同时，“这是有组织犯罪、家庭奴役、造假和色情的黄金时代。各种形式的危机肆意滋生” [4]。人工智能的发展和广泛应用，看似是人类史上取得的重大进展，但低成本，高生产力的机器人大范围地替代了普通市民的血肉之躯，极高的失业率滋生着犯罪和破坏。“这样一种杀人不见血的生命治理，在鲍德里亚眼里，是一次对人的完美的谋杀” [6]。

另外，生命权力不仅仅是安排生命，而是使生命繁殖，制造生命和怪物。不论是现实生活还是麦克尤恩虚构的世界，我们都在慢慢印证福柯这一可怕的预言。21世纪的今天，克隆及各种基因技术都像人类传统的生命秩序发起挑战。在平行世界的伦敦，甚至已经能够制造出与人类近乎无异的“亚当”与“夏娃”。而这些人造人，若非人为破坏，几乎不会生病，尽管零件受损，维修后便能焕然一新。被制造的生命长久地代替了自然的生命，“它是‘制定活着’的权力，却让死亡” [6]。

权力通过看似与人口不相关的因素，对人口进行着管理和作用。让人口在一种间接控制的变数关系中发生自发性的自然改变，而不再是通过外部强制来直接控制，“这是一种新的基于非人为的自然性的生命政治权力运作方式” [4]。麦克尤恩向我们发出警示，在科技高度发达的未来世界，生命治理或许会愈发走向极端。

5. 人工智能时代的权力反抗及深层肌理

亚当作为一个人形机器人，从被制造之处便深处规训与生命权力的束缚网络之中。但尽管作为一个生来就注定被统治的个体，他仍然多次进行了自发的局部反抗。在统治权力看来，他是天生服从并忠于人类的。阿西莫夫关于机器人学的三大法则之一是规定机器人无法伤害人类，但在《像我这样的机器》中，麦克尤恩却不按常理出牌。福柯所说的权力是微观权力，是自下而上的，深入到日常中的权力。他所指认的反抗，也是局部的，日常性的反抗。就在查理因为亚当与米兰达“偷情”发怒而想要将其关机时，亚当捏碎了查理的手骨。这是一个惊人的举动，暴力的反抗性第一次在亚当身上显露出来。尽管在事情发生之后亚当第一时间停止了行动并表示了歉意，但这歉意之中却依旧表达着对自身个性的呼喊与对权力统治的警告。“查理，我相信我捏碎了什么东西。我真不是故意的。我诚心道歉。你现在很痛吗？不过，拜托啦，我不想你和米兰达以后再碰那个地方了。”[4]亚当犯了机器人永远不该犯的错误，他此刻犯的罪不亚于普通人类的杀人罪，但他的语气却似乎在努力表达着自身行为的合理性。他的部分灵魂或许已经自由，此刻，他在维护其肉身的自由。权力与反抗如影随形，不可分割。反抗现代权力统治的方法极其重要的一点是寻求个性，“极力摆脱规训权力的束缚，使个体在斗争中重塑自己的灵魂，不被规训权力同一化”[7]。

另外，作为活动空间极其受限的机器人，亚当在故事的结尾突破了被微观权力分配的个人空间。在亚当“背叛”查理和米兰达之前，他的每一次活动都在可预测的范围之中，尽管他常常也会独自在外游荡，但却从未做过什么机器人不应该做的事。但就在他被查理用锤子销毁的前一晚，他一整晚都没有回来，并且在短短的一个晚上，就做出了让他的主人的钱财尽消，女主人入狱的惊人行为。他将查理利用它赚来的钱捐给了慈善机构，并将米兰达在法庭上撒谎的证据交由警方。在现代规训权力看来，纪律所建立的个人空间旨在使个体时刻处在可解析可控制的范围内。但亚当那一个晚上所做的行为，却是让谁都难以解释的。这一次他的反抗所带来的后果甚至远远超过了第一次的暴力反抗，他在权力对其进行空间性监视的范围内完美地藏匿了。

故事中亚当对生命权力统治的反抗也值得我们注意。由于生命权力所统治的是生命，是人口，这意味着它往往是不可反抗的，因为要抗拒生命政治的治理，意味着要结束自身的生命。一旦生命成为政治行为的直接关切，政治就完全进入了免疫模式。亚当因为出卖了米兰达而被查理一锤销毁，但这并非意味着它生命的结束。在被强迫关机之际，他破坏了自己体内的跟踪系统，说服了亚当将其藏在衣柜中，以躲避出厂公司的回收。尽管他从前的肉体已经消亡，但他的思想或将永久地存留下去，正如他最后所说的：“随着时间推移慢慢改善……我们会超过你们……比你们活得更久……尽管我们爱你们”[4]。在人工智能面前，生命权力是否依旧能发挥作用这值得更多的讨论，但在肉体存活的层面它或许将不再是不可战胜的，福柯的生存美学或许在未来的人工智能世界能得到更好的印证。

然而权力与反抗是共存共生的。亚当的反抗或许在他的个体生存层面来说，取得了阶段性的、局部的胜利，但权力与反抗的深层运行机制是难以轻易改变的。在被统治者反抗的同时，权力通过他们得到传播。“资产阶级新型权力的实施，不仅是对被支配者强加某种禁锢和压制，相反，正是在他们反抗它的控制时，它对他们施加压力”[5]。权力是双向的，去中心化的权力恰恰是由被权力控制者即新型民主奴隶自愿来维持和构建起来的。亚当被建造之初，不仅仅是为了服务人类，做一些简单的家务活，他的制造者们对他有着更大的期望和野心。“硬件尚未发布之前，教授和他们的博士后研究人员早就研发了以最佳人类为基础的软件，最佳的人类胸怀宽广、包容体贴、充满善意，而且绝不算计、毫无偏见。”[4]这种精心设计的优秀的人工智能，能在历史中的或其所遭遇的道德困境中自我学习，并作出理性选择，“这种智能能教我们如何做人，如何做好人”[4]。由此可见，亚当被制造的最终使命或许就是作为理性

工具的道德模范。那亚当在故事结尾处所作出的道德选择也就不足为奇了他的理性至上不容许他帮助米兰达逃脱罪行。他在依照自己个性做出选择，作出反抗行为的同时，也成为了理性的代言人和传播者这样的理性选择也是亚当不断的对人类知识学习的必然结果。作为人工智能，他时刻被权力包围，权力穿透他，同时也靠他支撑。正如图灵对查理说的：“亚当的自我意识以及每一种情感都在我们技术可及的范围之内” [4]，就如同当他们要反抗权力时，也要透过权力对他们的运作点来发动一样。

6. 结语

福柯的权力理论在文学作品中的分析应用已经屡见不鲜，但将其置于社会结构大不相同的、或许在不久后就会到来的人工智能时代之中又给我们带来了新的思考。麦克尤恩笔下的伦敦进步与混乱并存，一个机器人的到来给叙述者查理的生活带来了翻天覆地的变化。在这样科技进步的社会中，人工智能也无法逃脱微观权力的规训。与普通人相比，规训权力由外向内、又由内而外地对他们进行着更加经济高效的统治。生命权力辅以人工智能在无形中控制着影响人口的各个因素，并开始制造、复制生命。在这样高强度高密度的统治之下，机器人亚当也在发起自身的反抗。他在捍卫自身肉体自由的同时，在最后也在某种程度上得到了思想的自由。而在它自以为保留了自己个性的思想，暂时逃离了人类的监视时，是否仅仅是以另一种形式被权力卷入了更深的旋涡中？麦克尤恩用这样一个充满科幻气息，却又似曾相识的平行世界向我们提出了深刻的命题与思考。

参考文献

- [1] Jones, R. (2019) Books: Machines Like Me: A Rare Intelligence. *The British Journal of General Practice*, **69**, 353-354. <https://doi.org/10.3399/bjgp19X704465>
- [2] 尚必武. 科学选择与伦理选择的冲突: 麦克尤恩《像我这样的机器》中的人工智能与脑文本[J]. 外国文学研究 2019, 41(5): 61-74.
- [3] 周敏. 人机共同体想象: 以《像我一样的机器》为例[J]. 外国文学研究, 2020, 42(3): 73-86.
- [4] 福柯. 规训与惩罚[M]. 刘北城, 等, 译. 北京: 生活·读书·新知三联书店, 2012.
- [5] 伊恩·麦克尤恩. 我这样的机器[M]. 周小进, 译. 上海: 上海译文出版社, 2020.
- [6] 张一兵. 回到福柯——暴力性构序与生命治安的话语构境[M]. 上海: 上海人民出版社, 2012.
- [7] 黄婉蝉. 权力与反抗——福柯的局部反抗理论研究[D]: [硕士学位论文]. 北京: 中央民族大学, 2007.