

美德经济

人工智能领域中“伦理”的流通

编者：

Thao Phan、
Jake Goldenfein、
Declan Kuch、
Monique Mann

理论按需 |

网络文化研究所出版读本
系列 |

第46号

中文初译 v0.1

美德经济：人工智能领域中“伦理”的流通

INC “理论按需” 丛书第46号

美德经济：人工智能领域中“伦理”的流通

编者：Thao Phan、Jake Goldenfein、Declan Kuch、Monique Mann

文字编辑：Elena Gomez

封面设计：Katja van Stiphout

版式设计 with EPUB制作：Jasmin Leech

网络文化研究所出版，阿姆斯特丹，2022年

纸本ISBN：9789492302960；EPUB版ISBN：9789492302977

联系方式：网络文化研究所；阿姆斯特丹应用科学大学（HVA）；DMCI学院；Benno Premsele大楼；Rhijnspoorplein广场；04A07室；1019 GC 阿姆斯特丹；荷兰。电子邮箱：info@networkcultures.org；网站：www.networkcultures.org

订购纸本或免费下载本出版物：www.networkcultures.org/publications

订阅INC通讯：www.networkcultures.org/newsletter

本出版物采用知识共享“署名—非商业性使用—相同方式共享4.0”许可（CC BY-NC-SA 4.0；原书称Unported）。许可见原书所列 www.creativecommons.org/licenses/by-nc-sa/4.0/；可用规范地址：<https://creativecommons.org/licenses/by-nc-sa/4.0/>。

网络文化研究所（Institute of Network Cultures）

中文初译制作说明

中文初译 v0.1 | 2026年10月13日

本稿从2022年英文版直接翻译，完整收录封面、版权页、目录、导论、三部分文章与访谈、原注、参考文献、致谢、撰稿人简介及原封底。原版共173个PDF物理页；两处空白页不机械重印。中文正文适配原开本与无衬线版式，保留原书色彩、封面线条与分部留白。封面与取得的原封底均作中文适配。

各章均由未参与该章初译的AI代理逐段对读，尚未经独立人工终审；“中文初译 v0.1”不代表出版社正式出版、认可或中文首译。人物简介、职务及文内时态均按2022年底本保留。

原书页下注完整移至各章末“原注”，保留各章原注编号与双向跳转。中文PDF目录使用译本实际页码。可编辑稿中的非空白源页标记定位到包含该页内容的首个译段；跨页自然段共用锚点，原注另见章末，空白源页只保留位置标记。完整的源页与译段关系见交付目录中的源页映射。参考文献及原注中的文献著录保留可检索的原文信息；有题名者附中文译名。

图1的抗议标语作AI中文化重绘，原照片另存供核对；原有导弹删线、位置关系和图片署名保留。封面采用原PDF装饰曲线的矢量转写，文字重新排版。原版权页误植的书名Economics与编者姓Goldfein，按封面、封底及出版社书目统一为Economies与Goldenfein。其他已识别差异在译注及本书校勘记录中说明。

原书采用CC BY-NC-SA 4.0许可。本地译稿按相同许可保留署名、非商业与相同方式共享条件；第三方转载说明仍随相应章节保留。完整来源、图像处理记录及验收报告见本书交付目录。

目录

导论：美德经济 Thao Phan、Jake Goldenfein、Declan Kuch和 Monique Mann	6
第一部分：主体	25
一分钱买你的思想？资本、共谋与AI伦理 Corinne Cath和Os Keyes	26
攫取主义伦理 Sarah Pink	43
“伦理AI”的发明：大型科技企业如何操纵学术界以规避监管 [转载] Rodrigo Ochigame	53
第二部分：场域	61
生态灭绝并不合乎伦理：政治生态学与资本主义AI伦理 Sy Taffel、Laura Bedford 与 Monique Mann	62
日常AI伦理：通过人脸识别从全球走向地方 Angela Daly	95
以数据为宴：伦理、价值与风险偏好的计算 Tsvetelina Hristova 和 Liam Magee	121
军事美德与AI研究中“伦理”的限度 Michael Richardson	142
第三部分：行动	158
“公开的秘密”：Meredith Whittaker 访谈 访谈：Jathan Sadowski、Thao Phan	159
“退出者”：Lilly Irani、Alex Hanna、Khadijah Abdurahman 访谈 访谈：Jake Goldenfein	171
致谢	186
撰稿人简介	187

导论：美德经济

Thao Phan、Jake Goldenfein、Declan Kuch和Monique Mann

“AI伦理”做了什么？它为谁服务？它的目的是什么？如果我们问那些关心技术与社会、政治和环境正义问题的人，那些与受到自动化决策影响的社群一道开展实地调查的人，或在监管机构内部工作的人，预期得到的答案可能是：“什么也没做。”曾界定“科技反弹”的迅猛批评浪潮，如今正被一股更大的浪潮吞没，Elettra Bietti称之为“伦理抨击”。^[1]如果说科技反弹的核心，是揭露世界上一些最大的科技公司的不良行为，那么接下来的这股浪潮所针对的，则是“伦理”作为一种新产业议程的兴起，聚焦于成立AI伦理委员会、雇用AI伦理团队、资助AI伦理研究等行动。批评者有理由把这些行动理解为：试图将伦理工具化，^[2]把它降格为另一种产业资本，^[3]或通过收编与俘获研究者，来控制公共叙事。^[4]

许多批判性论述试图为相关领域带来透明与问责，而这些领域本身恰恰关注AI系统的透明与问责问题。AI伦理会议——最显著的是美国计算机协会每年举办的公平、问责与透明会议（FAccT）——开始设置完整的专题单元，将研究目光转向组织自身及其价值、治理过程和赞助政策，这标志着这些社群发生了重大的反思性转向。^[5]这些讨论面向的听众，很多本来就在AI伦理这一宽泛范畴之下工作。尽管如今越来越多的学者和行动者试图否认、乃至挑战“AI”和“伦理”这两个词，但对于学术界之外的人——那些并非整天泡在Twitter上、阅读Medium、Wired或《麻省理工科技评论》文章的人——而言，“AI伦理”大致就是公众眼中我们的工作。如果我们把伦理理解为“一种内化的、怀抱愿景的探究方式，旨在创造更好的世界”和“更公正的社会”，^[6]那么许多人不会反对这个称谓。反对意见源自伦理已经变成了什么。简言之，就是它被用来做什么、它服务于谁，以及它实际实现、而非应当实现的目的。

面对“AI伦理做了什么”这一问题，实际答案很少真的是“什么也没做”。日益增长的批判性文献表明，AI伦理中的“伦理”比预想的更可塑、更柔韧，也更易于被策略性地运用。据说，伦理既可以被用来洗白，^[7]也可以遭到抨击。^[8]它可以成为橡皮图章，^[9]或一个空洞姿态。^[10]它可以是遮羞布，^[11]是诱人偏离方向的岔路，^[12]或者，正如Lily Hu在尖锐分析中所写的，它可以是一场表演性的戏剧：

“硅谷用长长的金钱丝线操纵着科技伦理这具木偶；‘伦理’是一场表演，他们对此心知肚明。”^[13]简单说，大型科技企业所部署的伦理，比无所作为更糟。它充满分裂与争议，而且往往使令人担忧的“一切照旧”成为可能。在最坏的情况下，AI伦理不仅无用，而且“危险：它囤积了本该用于更有效工作的专业知识与资金”。^[14]因此，Karen Hao等知名科技记者^[15]曾严厉批评AI伦理社群：

空谈终究只是空谈——这远远不够。尽管许多组织对这些问题说尽了漂亮话，它们的AI伦理准则仍然含糊不清、难以实施。很少有公司能够展示，它们在评估和批准AI产品与服务的方式上作出了切实改变。我们正在掉进伦理洗白的陷阱，真正的行动被表面的承诺所取代。

本书从这里出发。最初征稿时，我们邀请学者和行动者回应日益盛行的AI伦理抨击，特别关注支撑AI伦理的财务结构，以及这些结构对从事这项工作的人的影响。我们——本书的编者——提出了这样的问题：是谁在资助有关AI技术伦理的研究？这些资助安排可能怎样巩固或加剧权力层级？由产业赞助的伦理议程，又如何影响关于AI系统的知识生产？

我们在征稿启事中指出，随着倡议者、产业内部人士和公民社会行动者对“不良AI”提出关切，如今已出现一波由社会科学引领的AI伦理研究。我们加入了日益壮大的讨论，也表达了一种担忧：随着这个研究与实践领域不断发展，它本身也面临一场伦理危机；这场危机使该领域研究者复杂的位置性凸显出来。最初迫使企业以更合乎伦理的方式行事的社会与政治压力，后来转化为对研究文化的各种渗透，形式包括会议赞助、正式研究伙伴关系、产学研联合聘任等等。^[16]

当然，这股产业资金的涌入，是一系列更广泛问题的症状：高等教育政治化；针对所谓“左翼知识分子倾向”的敌意文化被培植起来；以及新自由主义的连续两击——世界各地公共机构经费遭到撤回，同时获取外部产业资金的激励不断增强。例如，在澳大利亚，一个失灵的资助体系意味着：1) 少数现有的政府资助来源所提供的资金少于十年前，而且主要只有已经富裕的机构能够获得；2) 项目资助的结构主要在于激励那些被认为“可商业化”或对产业有价值的研究；3) 即便研究者通过了竞争极其激烈的遴选过程，^[17]项目能否成功仍可能受到政治干预。^[18]公立高等教育的公共资金逐渐撤回，意味着大学必须转向其他收入形式来求生。仍以澳大利亚为例：大学如今把教育当作次要事业，真正的业务是在管理投资组合之外从事其他牟利活动，并聘请前企业高管，像管理大型公司那样管理大学。结果是，大学只在教学与研究能够转化为影响国际排名的指标、进而兑现为国际学生学费的限度内，才重视它们。与此同时，财务纪律的压力层层下传，迫使大学尽可能多地榨取劳动者，甚至系统性地克扣员工工资。^[19]

这些因素把AI伦理置于尴尬境地：在财务上依赖于它们试图追究责任的组织。Hu指出：“一个令人不安的事实是，无论外部顾问委员会和大学多么宣称自己是‘第三方’，科技伦理机构实际上都寄生于霸权性科技产业持续不断的道德失误与失望之上。”^[20]在Hu看来，这些关系相互强化：企业利用AI伦理掩饰要求改变的政治诉求，而科技倡议团体则利用产业合作所带来的影响力与资源。^[21]科学技术研究学者Lee Vinsel把这种现象称为“批评式炒作”（criti-hype）——一种围绕新兴技术、既“助长炒作又以炒作为食”的自利性批评。^[22]

实际上,本书本身也始于一笔面向澳大利亚青年研究者的小额科研资助(9,000澳元)。这笔“工作坊项目资助”来自澳大利亚社会科学院,该组织主要依靠澳大利亚联邦政府拨款和院士会费支持。^[23]我们自己也处于职业生涯的早期或中期,机构对获取经费的期待,恰好与一个认识相遇:有关AI、尤其AI伦理维度的话题,非常容易获得资助。我们始终真诚地致力于探究资助、劳动与AI伦理这一主题,但对于资源和机构支持的实际需求,也从未远离。

作为人文与社会科学研究者,我们是幸运的:我们想回答的问题,不需要昂贵的硬件、算力、大型数据集访问权限或服务器空间;这些需求往往迫使STEM院系的许多同事作出无奈选择。我们的需求更基础:时间。既包括我们自己的时间,也包括在这个项目的整个生命周期里与我们交流的各位对话者的时间。大多数情况下,经费不仅能够保障资源,也可以积累影响力与权威,为获得资助的人赋予正当性和荣誉。有了经费,我们可以支付餐饮、文字编辑、行政和研究助理的费用,举办工作坊并推进出版物。但经费最宝贵的形式,是赋予我们必要的声望,使我们能够请他人为项目捐出时间,也允许自己从机构那里夺回自己的时间。

我们在这里较详细地说明这一点,是因为恰恰是这些平凡、实际的协商与机构要求,决定了AI伦理工作者的劳动条件。普遍的不稳定、不安全,以及被正常化的疲惫与过劳文化,使这类反思性和批判性对话很少能够发生,除非它们也能被调动起来,服务于战略目标与个人绩效指标。人文与社会科学等领域尤其能够感受到这些压力:它们承受院系经费削减,而且数十年来一直在对抗保守派主导的文化战争。^[24]在这种背景下,许多人转向AI伦理等有利可图的话题,或转向产业资助,以求生存,也就不足为奇。

正是在大学内部这些令人沮丧的条件下,大型科技企业成功地做到了Fred Moten与Stefano Harney那句名言所说的:“把反叛者变成国家的代理人。”^[25]“AI伦理学家”如今成为一个正当的职位名称,恰好说明了Moten与Harney的论点。曾经的批评者现在靠教授有关AI的社会、政治与伦理维度的专业课程和学位项目领取薪水,培养新一代专业技术官僚,据称他们能比上一代更好地开展AI工作。正如Ben Tarnoff敏锐地指出的,科技反弹的一个意外效果,是“为一个新的专家阶层举行了一场大规模资格认证活动;‘AI伦理’、‘负责任创新’及类似事业吸引了大量资金与关注”。^[26]这些专业化与资格认证形式,只是在为那些曾经的批评对象——企业行动者——的利益服务。套用Moten与Harney的话:AI伦理对于大型科技企业的企业议程而言,“不只是盟友”,而是“使这一议程完成的尝试”。^[27]

下面几节将更明确地讨论本书的主题与概念。我们首先简要回顾AI伦理的兴起,然后讨论我们所看到的当代现象:伦理被物化为商品形式。接着,我们逐章介绍,说明各章如何展示那些界定了我们所谓“美德经济”的庞大而多样的流通网络,最后简要讨论AI伦理此后的发展路径。

AI伦理的兴起

“科技反弹”（tech-lash）一词在2013年进入主流语汇，当时《经济学人》的Adrian Wooldridge提出，公众情绪正在转向反对硅谷科技精英。^[28] 产业对物质不平等的加剧、Google Glass等失败且明显不时髦的消费产品、反复发生的隐私失守、不合理的税收优惠，以及问题重重的政治立场，动摇了公众对“大型科技企业”的看法。^[29] 例如，Google“组织全世界的信息”的努力，^[30] 被重新理解为蕴藏着造成深刻社会伤害的能力。分析者意识到，那些曾被视为“开放互联网”等解放性话语旗手的公司，如今已经把对在线服务和数据流的控制，转化为令人担忧的牟利与支配形式。硅谷的解放承诺远没有促成与既有社会经济状况的革命性决裂，而是被揭露为产业资本主义的又一种表现。在学术界，批判性学者早已描述过硅谷反文化的自我呈现与其对资本的不懈追逐之间的矛盾。^[31] 但AI伦理作为一个研究领域，在科技反弹的同时、也作为其一部分获得正当性，激励了更多来自社会科学和技术学科的知名学者，对科技产业的服务与结构作出新的批判性分析。

大型科技企业的一项主要产品——机器学习（ML）——处在这场重新评估的最前沿。人们越来越认识到，它经常产生带有偏见的结果，深刻影响边缘化人群的生活。2013年，主流新闻开始报道：机器学习技术因为接受的是历史数据——而这些数据在历史上就存在偏见——的训练，如何再生产性别化和种族化的歧视。^[32] 人们也越来越意识到，算法评分系统在就业、社会服务与警务等领域犯下的错误、误识别和有问题的群体画像。^[33]

由于具有明确的伤害能力，又牵涉巨大的商业利益，机器学习成为AI伦理这一领域围绕其定位的一组技术。例如，Cynthia Dwork等人早期关于机器学习分类公平性的技术工作——其中包括FATML会议的早期组织者Moritz Hardt——^[34] 借鉴H. Peyton Young、John Roemer和John Rawls政治理论中的平等与公平概念，探索了进行“公平”分类的方法。一旦这些技术问题的政治和社会意义被揭示出来，公平与偏见便迅速成为人文与社会科学的研究对象。

2013年3月，哈佛大学数据隐私实验室的创始人兼主任Latanya Sweeney发表了一篇有关在线广告投放中种族歧视的影响深远的论文，将批判的目光引向商业平台如何通过其广告投放系统再生产种族歧视模式。^[35] 不久后，时任微软研究院首席研究员的Kate Crawford在《哈佛商业评论》发表了《大数据中的隐蔽偏见》。^[36] Crawford将修正（大）数据科学界定为研究技术与社会关系的下一个前沿，建议技术人员与社会科学研究者对接，并建议计算科学认真面对定性方法。在《大西洋月刊》上，Nicholas Diakopoulos把偏见问题与机器学习的不透明问题联系起来。^[37] 此外，Safiya Noble通过揭示搜索引擎中种族与性别偏见的研究，将许多涉及数据驱动系统、商业利益与种族偏见的问题汇合在一起。^[38] 2014年8月，Solon Barocas与Andrew Selbst公开了他们的文章《大数据的差别影响》，^[39] 细

致说明大数据技术如何产生在法律上具有意义的歧视性待遇。〔译注：原文此处将Diakopoulos拼作Diakopoulos，保留原文。〕

2014年前后，公民社会也开始关注与大数据和AI有关的社会正义问题。例如，民权与人权领导会议于2014年2月发布了《大数据时代的民权原则》，^[40]提出应通过问责与公平来应对嵌入其中的歧视。Upturn等组织使这一议程更加具体，^[41]将数据驱动决策的潜在伤害与更广泛的社会正义问题联系起来。不久之后，奥巴马白宫（总统行政办公室）发布了2014年大数据报告，将大数据中的偏见与歧视识别为可能造成重大物质与非物质伤害的风险。

有关算法偏见与机器学习公平性的关切，也与算法问责方面的相关学术兴趣联系起来。纽约大学在2013年5月举办了“治理算法”会议，提出如何把“算法问题转化为富有成效的研究对象”。^[42]Data & Society也于2014年3月提出了围绕算法问责的研究议程。^[43]2015年2月，纽约大学在2013年活动的基础上举办了“算法与问责”会议，明确将主题界定为算法权力在透明、公平和平等待遇方面带来的挑战，并邀请主要来自人文与社会科学领域的演讲者。一个更明确面向技术的论坛，也在2014年的FATML工作坊中形成；它希望探索“如何以计算上严谨的方法描述并处理这些问题”，^[44]并在一个复杂、跨学科的问题界定之中，为大体属于技术性的解决方案提供交流场所。该工作坊持续举办至2018年；同年又增设了规模更大、更具跨学科性的FAT*会议，汇合社会科学与技术领域。该会议于2019年成为ACM旗下会议，并在2020年更名为FAccT。如今，已经有众多专门的AI伦理工作坊、专业组织和会议，有些独立举办，有些是更广泛技术活动的一部分，研究算法偏见、技术客观性的虚假性，以及以问责、透明和公平为形式的解决方案。

虽然AI伦理是作为对产业失误的回应而兴起的，但从这一领域的早期开始，大型科技企业就长期在支持它方面扮演核心角色，并持续通过大量资金与支持参与学术环境。毫无疑问，产业对AI伦理研究的支持，促成了大量高质量的独立学术成果。但与此同时，这也提出了围绕研究者位置性以及界定研究“价值”的机构动态的复杂问题。

考虑到AI伦理分析的核心对象本就是商业应用与工具，产业与学术界围绕这些特定问题汇聚起来，并不令人意外。然而，恰恰是利益与价值观明显不同的利益相关者网络的发展，使这里被称为“美德经济”的新型流通成为可能。^[45]Lily Hu谈到AI伦理领域的旗舰会议时指出：“总体而言，FAccT研究者并不是在对着虚空呐喊；恰恰相反，许多人实际上会在会后由企业赞助的鸡尾酒会上见面，讨论跨机构、跨利益的合作。”^[46]确实，FAccT这样的活动本质上充当了学术界与产业界之间的接口，提供一个平台，让这些群体及其相互竞争的利益得以整合。它们创造社会与职业网络，既可能促成伙伴关系与合作，也可能成为直接进入产业就业的通道。这些旗舰活动的规模，也意味着它们依靠大型科技企业的经费来补贴举办活动的后勤

成本。这在专业学会与它们试图问责的组织之间，形成了特定的义务；这些义务确保大门始终向产业利益的流入敞开。

这些社会与智识交流，为研究、研究者及声誉被物化为商品形式奠定了基础，使它们能够在产业机构与学术机构之间、并穿行于其中流通。这些流通采取多种形式，为所有参与者吸纳、生产和交换明确的财务价值以及声誉价值。例如，学术伦理成果可能成为获取产业资助、产业支持的学术平台或产业职位的杠杆。与此同时，大学赞许这些发展，利用外部资金、声望以及与产业互动相关的“影响力”的流入。而大型科技企业则始终受益于自身伦理资质与伦理事业得到强化和正当化。

这类流通包含一个矛盾：高价值的伦理商品，唯有通过有限的“伦理效应”才能获得并维持价值。换言之，为了能够在学术界与产业之间自由流通，其伦理内容——就其对世界、或至少对它所指向的企业的影响而言——必须被抹除，或至少受到限制。著名AI伦理研究者Timnit Gebru遭其前雇主Google恶劣对待的事件表明，^[47]当伦理商品从根本上挑战商业模式时，用来交换它的职位与地位就会消失，甚至还会发生更糟的事。产业所获得的伦理商品，其使用价值或许表现在别处：不在于引导产业与技术实践的实质性变革，而在于构筑声誉大厦，使之能够：1) 保护商业活动免受可能影响利润的结构性批评；2) 鼓励AI伦理领域朝对产业友好的方向发展。伦理可能采取的多种形式、其流通渠道、它所做或未做的工作，以及它的要求如何与研究者的经验相交，正是本书各篇文章探索的主题。

本书各篇文章

本书是对伦理被物化为商品形式这一现象的集体回应。它探索产业参与伦理AI研究如何创造了新的美德经济：一个庞大的行动者网络，分布于产业、公民社会和大学，以服务与产品的形式生产和流通伦理。作者们既带来了批判性视角，也带来了在这种经济中生活所面对的挑战、困境与机遇的亲身经验。他们的经历多样，来自法律、人类学、犯罪学、媒介与传播研究、科学技术研究、政治经济学等一系列学科背景。有些作者是资深学者，是在该领域积累了数十年经验的教授；另一些则刚刚开始职业生涯，在AI伦理资助狂潮最盛时进入产业。不过，他们共同关心技术、权力与社会正义问题；尽管很少有人——如果有的话——会称自己为“AI伦理学家”，他们仍无不与该领域发展过程中出现的争议和辩论紧密纠缠。

在以下各章中，这些作者对商品化与抵抗的战术和策略发声并作见证。每一章都探讨这些动态如何在不同场所与地带展开。把这些故事放在一起，我们就开始看到行动者、机构和利益之间共同的轨迹与流动。一个令人悲哀的讽刺是：伦理越是作为商品流通，它所能做的伦理工作就越少。然而，正如前文讨论的、以及各章将要展示的，即使采取商品形式，AI伦理仍然总能被投入工作，去做某件事、为某些人的利益服务。

本书分为三个部分：主体、场所与行动。第一部分的作者从自身主体性与主观经验出发，叙述这些复杂资助安排将劳动者置于其中的矛盾与困境。

在《一分钱买你的思想？资本、共谋与AI伦理》中，Corinne Cath和Os Keyes把产业对博士奖学金与合作项目的赞助，描述为一种精明的投资。他们通过在美国与英国攻读博士学位的经历呈现片段，提供生动的细节，展示当自己的工资与自己的敌人捆绑在一起时，一线工作者如何应对、协商随之产生的紧张与不适。“批评并不能避免共谋，”他们写道；正如这些例子所展示的，在许多情况下，批评本身成了一种收编方式。他们以呼吁女性主义拒绝作结：通过承认任何参与——无论多么批判——都不可能处于权力之外，这一立场有助于把研究者放在权力关系之中加以理解。

在《攫取主义伦理》中，Sarah Pink把伦理描述为“诱饵，借以从公众或用户那里提取对技术的信任”。她认为，技术解决主义的设计方法使生产者与普通人的日常生活脱节，进而培育出一种工具主义伦理观。人类伦理价值只有在可被用来制造“伦理机器”、而这些机器又可被展示以产生信任的限度内，才有其用途。她呼吁回归日常伦理；日常伦理本质上难以捉摸且不稳定，因此不那么容易被俘获、成为投资对象——至少无法以产业行动者所设想的方式做到。

在本部分最后一篇文章中，Rodrigo Ochigame描述了伦理如何被界定为一种“友善的批评”，可以充当“建立商业关系的杠杆”。和Cath、Keyes一样，他叙述了自己作为研究生研究者，在麻省理工学院媒体实验室AI伦理小组工作的时光。他坦率评论了笼罩该实验室及其前主任Joichi Ito的伦理丑闻，毫不回避地评估媒体实验室在维系硅谷议程方面的作用。Ochigame的文章最初发表于调查新闻网站*The Intercept*，是公开反对那些多数人只能从窃窃私语中听说的肮脏运作机制的出色范例。我们很荣幸获得他的许可，在此重刊。

本书第二部分“场所”转向具体情境中的案例研究，理解大型科技企业如何在AI应用的特定领域运用“伦理”实践。

在《生态灭绝并不合乎伦理：政治生态学与资本主义AI伦理》中，Sy Taffel、Laura Bedford和Monique Mann把AI伦理描述为一种企业手段：借此界定伦理实践，使其避开人为造成的地球性伤害。他们在资源开采地、数据中心和电子垃圾处理场之间穿梭，详细讨论AI与机器学习的社会和生态影响。这一章对“绿色AI”的幻想如何维持不公正、不合伦理、而且明确造成生态灭绝的企业实践，提出了严厉批评。

在Angela Daly看来，推动AI伦理的企业议程，使它变得抽象、脱离现实且去政治化。和Sarah Pink一样，她也呼吁在分析上回到日常生活的场所，理解一线人群复杂的生活现实，并凸显抽象伦理原则的不足。Daly以人脸识别为具体例子，离开高

谈阔论的伦理原则世界，转而关注一个通过抗议、异议与组织行动而付诸实践的伦理协商世界。

在讨论全球标准与标准制定的章节中，Tsvetelina Hristova和Liam Magee说明，伦理如何被用作将风险社会化的载体，从而使其扩大规模并转化为一种价值。他们的研究拓展了通常与美德经济相联系的场所和实践，转向ISO（国际标准化组织）框架与分委员会这些高度官僚化的领域。他们在此考察中等强国如何把标准制定当作经济和政治策略，而其效果就是把伦理转化为字面意义上可出口的商品。

最后，Michael Richardson以一章关于研究伦理悖论的典型场所——国防与军事AI——的文章结束本部分。他描述了对伦理的关注如何使与国防部门合作的研究者免于面对致命暴力所带来的不安问题，同时又为使用自主武器加强打击效果与伤害提供正当理由。他写道：“国防研究者和企业不仅能够成为有美德者，还能使战争也具有美德。”他强调，国家资助减少如何迫使大学通过企业与军事伙伴关系，拓宽收入来源。他提醒我们：“大学毕竟首先是帝国与殖民的机构，甚至比它们作为学习、知识生产和异议场所的身份更甚。”

第三部分也是最后一部分“行动”，以对知名学者和行动者的两场访谈结束本书，反思那些直接反抗产业资金与影响力政治的行动时刻。通过讨论，我们得以一窥其他可能性、行动主义可能实现什么，以及学者需要怎样的反思性，才能理解塑造其工作生活与主体性的复杂力量和要求。在Meredith Whittaker、Jathan Sadowski和Thao Phan参与的《公开的秘密》，以及Lilly Irani、Alex Hanna、J. Khadijah Abdurahman和Jake Goldenfein参与的《退出者》中，这些在美德经济不同位置上拥有丰富经验的学者，描述机构劫持我们作为研究者的自我意识的不同方式，以及我们用来（重新）想象和／或（重新）定义大学的努力，或许更应投入到像任何其他工业企业一样，以大学劳动者的身份组织我们自己。

这些学者凭借跨部门的丰富多样经历，描述了资助在社群、同事情谊以及支持人们挣得工资这一背景下的复杂性，也描述了这些东西如何被利用为内化纪律的工具。他们对产业影响学术界的内部机制、以及联盟如何围绕资助优先事项形成，有着切身理解；他们说明这些机构如何鼓励某些类型的批判性工作，同时破坏可能削弱产业利益或资助关系的激进努力。他们提出的替代方案，是把工作场所组织化作为工具，使学者摆脱那些服务于掌控我们者利益的声望要求，并由此使学者在美德经济中所从事的工作去商品化。

结语：从AI伦理到美德经济

本书属于AI伦理／AI与社会责任领域日益增长的一组文献，这些文献试图提出“关于知识创造条件的自我反思性批评”。^[48]这种批评方式建立在一个悠久传统之上：研究者把自己的实践、社群与机构作为批判分析和反思的对象。批判种族研究、原住民研究、黑人女性主义理论以及其他女性主义学术分支，都提供了这种方式的典范。^[49]虽然这类工作的基调往往是批判性的，但其意图是促成积极变化，并要求组织与研究社群对其宣称的标准和价值负责。近年来，许多学者把注意力转向产业与学术界之间不断扩大的接口，以及它对伦理主题的影响——伦理既是一个研究领域（如生命伦理、AI伦理），也是管理知识创造的一套指导原则（如多样性、公平与包容计划）。^[50]

本书力图凸显研究者在这一领域中位置性的复杂性，揭示我们所有人在选择研究课题、方法和合作伙伴时面对的困境。许多作者直接在产业、产业的研究部门，或接受产业资金及其他支持的大学院系与研究机构中工作过。^[51]多篇文章从个人经验出发，深入揭示大型科技企业的制度与组织安排，并为塑造这些情境中研究实践的劳动条件作见证。我们无意因为同事在资本主义下被迫作出的选择而点名指责他们。我们关心的是，产业如何既建立又利用这一研究环境中的激励结构。我们希望，这能促进大学劳动者进一步反思和分析自己在该领域的工作条件。

在学术情境中，研究与研究者的评估机制，是知识创造的核心。^[52]这历来是一个充满争议的领域，而且大型科技企业在其中的地位日益重要。^[53]在策划与制作本书时，我们采取了集体编辑模式，将其他章节分配给各章的（主要）作者审阅。我们这样做，是为了实施一种建设性的“同行评审”方法，抑制学科把关，并容纳这部文集中贯穿的具有挑战性的议题和主题。我们的目标，是鼓励审阅者对自己的反馈负责，同时紧密、建设性地共同工作，提高文字与论证的质量。我们也希望参与某些机制，使学术共同体能够加强对自身工作条件的控制。

对我们这些编者而言，这是一项建立团结的实践，与匿名同行评审过程相去甚远：商业出版商通过后者榨取学术劳动，而大学的激励结构通常并不认可这种劳动，劳动者也普遍感到它不受重视且令人不快。同样的愿望——（重新）想象与（重新）定义大学工作经验——也影响了我们对出版社的选择：网络文化研究所（Institute of Network Cultures, INC）。该出版社在跨学科知识生产与参与方面有着历史与声誉，通过开放获取出版与倡议，回应与数字网络有关的紧迫问题。

但在一个由以下各章所描述的动态界定的研究生态中，这些当然只是边缘性的行动。如何找到前路、穿行于美德经济之中，是一项艰巨挑战。本书这些文章，是为科技俘获、收编与妥协的动态命名这一过程的一部分。我们希望学者、倡议者、行动者和政策制定者，能够吸收这些对知识创造与传播条件的反思性批评，也认识到知识工作者面对的妥协与权衡——有利益诉求的机构掌握着对他们的权力。考虑到

塑造学术声望网络的同僚亲近关系政治，这必然令人不适。然而，只有命名这些动态，才能处理它们，才能提出那些使我们得以想象并要求另类未来的问题。

资助披露

本项工作获得澳大利亚研究理事会“自动化决策与社会卓越研究中心”支持[资助编号CE200100005]。我们尤其感谢该中心的墨尔本法学院节点，为本书完成提供财务支持。Declan Kuch感谢西悉尼大学校长研究员计划对印刷给予的支持。本书建立在2021年7月8-9日于迪肯大学举办的“美德经济：AI与数字文化中‘伦理’的流通”工作坊基础之上；该工作坊由澳大利亚社会科学院2021年工作坊项目资助。

参考文献

Abdurahman, J. Khadijah. ‘On the Moral Collapse of AI Ethics’ *Medium* blog, 8 December 2020, <https://upfromthecracks.medium.com/on-the-moral-collapse-of-ai-ethics-791cbc7df872>. [题名中译：论AI伦理的道德崩塌。]

Abdurahman, J. Khadijah. ‘Fired for Speaking Out Against Anti-Blackness and Ethnic Cleansing: Open Letter to Cornell Tech’, *Medium* blog, 3 June, 2021, <https://upfromthecracks.medium.com/fired-for-speaking-out-against-anti-blackness-and-ethnic-cleansing-open-letter-to-cornell-tech-3b3f669cf69f>. [题名中译：因公开反对反黑入主义与族群清洗而被解雇：致康奈尔科技学院的公开信。]

———. ‘On the Moral Collapse of AI Ethics’, *Medium* blog, 7 December, 2020, <https://upfromthecracks.medium.com/on-the-moral-collapse-of-ai-ethics-791cbc7df872>. [题名中译：论AI伦理的道德崩塌。]

Ahmed, Sara. *On Being Included: Racism and Diversity in Institutional Life*. Durham: Duke University Press, 2012. [题名中译：论被纳入：制度生活中的种族主义与多样性。]

———. ‘The Language of Diversity’, *Ethnic and Racial Studies* 30 (2, 2007): 235–56. [题名中译：多样性的语言。]

Australian Research Council. ‘Selection Report: Discovery Early Career Research Award 2022 | Australian Research Council.’ Australian Government | Australian Research Council, 2022. <https://www.arc.gov.au/funding-research/funding-outcome/selection-outcome-reports/selection-report-discovery-early-career-research-award-2022>. [题名中译：遴选报告：2022年发现项目青年研究者奖 | 澳大利亚研究理事会。]

Barnes, Joel. ‘Defunding Arts Degrees Is the Latest Battle in a 40-Year Culture War’, *The Conversation*, 3 July 2020, <http://theconversation.com/defunding-arts-degrees-is-the-latest-battle-in-a-40-year-culture-war-141689>. [题名中译：撤销文科学位资助是一场持续40年文化战争的最新战役。]

Barocas, Solon, and Andrew D. Selbst. “Big Data’s Disparate Impact.” *California Law Review* 104 (3, 2016): 671–732. [题名中译：大数据的差别影响。]

- Benjamin, Ruha. *Race After Technology: Abolitionist Tools for the New Jim Code*. 1 edition. Medford, MA: Polity, 2019, <https://ebookcentral-proquest-com.ezproxy-b.deakin.edu.au/lib/deakin/detail.action?docID=5820427>. [题名中译: 技术之后的种族: 废除新吉姆代码的工具。]
- Bietti, Elettra. ‘From Ethics Washing to Ethics Bashing: A View on Tech Ethics from Within Moral Philosophy’, *SSRN Electronic Journal*, 2021. [题名中译: 从伦理洗白到伦理抨击: 从道德哲学内部看科技伦理。]
- Birhane, Abeba, Ruane, Elayne, Laurent, Thomas, Brown, Matthew S., Flowers, Johnathan, Ventresque, Anthony, and Dancy, Christopher L. ‘The Forgotten Margins of AI Ethics’ in *2022 ACM Conference on Fairness, Accountability, and Transparency*, 948–58. FAccT ’22. New York, NY, USA: Association for Computing Machinery, 2022. [题名中译: AI伦理被遗忘的边缘; 2022年ACM公平、问责与透明会议。]
- Bourdieu, Pierre. *Science of Science and Reflexivity*, trans. Richard Nice, Chicago: University of Chicago Press, 2004. [题名中译: 科学之科学与反思性。]
- Burn-Murdoch, John. ‘The Problem with Algorithms: Magnifying Misbehaviour’, *Guardian*, 14 August 2013, <https://www.theguardian.com/news/datablog/2013/aug/14/problem-with-algorithms-magnifying-misbehaviour>. [题名中译: 算法的问题: 放大不当行为。]
- Crawford, Kate. ‘The Hidden Biases in Big Data’, *Harvard Business Review*, 1 April 2013, <https://hbr.org/2013/04/the-hidden-biases-in-big-data>. [题名中译: 大数据中的隐蔽偏见。]
- Diakopoulos, Nicholas. ‘Rage Against the Algorithms’, *The Atlantic*, 3 October 2013, <https://www.theatlantic.com/technology/archive/2013/10/rage-against-the-algorithms/280255/>. [题名中译: 对算法的愤怒。]
- Dormehl, Luke. ‘Algorithms Are Great and All, But They Can Also Ruin Lives.’ 访问日期: 2022年10月14日. <https://www.wired.com/2014/11/algorithms-great-can-also-ruin-lives/>. [题名中译: 算法固然出色, 但也可能毁掉人生。]
- Dwork, Cynthia, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. ‘Fairness through Awareness.’ In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference on - ITCS’ 12*, 214–26. Cambridge, Massachusetts: ACM Press, 2012. [题名中译: 通过觉察实现公平。]
- Elliott, Carl. *White Coat, Black Hat: Adventures on the Dark Side of Medicine*, Boston: Beacon Press, 2010. [题名中译: 白大褂, 黑帽子: 医学阴暗面的历险。]
- . ‘Why Clinical Ethicists Are Not Activists’, *Hastings Center Report* 51 (4, 2021): 36–7. [题名中译: 为什么临床伦理学家不是行动者。]
- Gansky, Ben, and Sean McDonald. ‘CounterFAccTual: How FAccT Undermines Its Organizing Principles’ in *2022 ACM Conference on Fairness, Accountability, and Transparency*, 1982–92. Seoul Republic of Korea: ACM, 2022. [题名中译: CounterFAccTual: FAccT如何破坏自身组织原则; 2022年ACM公平、问责与透明会议。]

- Gebru, Timnit. 'For Truly Ethical AI, Its Research Must Be Independent from Big Tech', *Guardian*, 6 December 2021, <https://www.theguardian.com/commentisfree/2021/dec/06/google-silicon-valley-ai-timnit-gebru>. [题名中译: 要有真正合乎伦理的AI, 其研究必须独立于大型科技企业。]
- Goldenfein, Jake, and Daniel Griffin. 'Google Scholar – Platforming the Scholarly Economy', *Internet Policy Review* 11, no. 3, 29 September 2022, <https://policyreview.info/articles/analysis/google-scholar-platforming-scholarly-economy>. [题名中译: Google Scholar: 学术经济的平台化。]
- Hao, Karen. 'A Leading AI Ethics Researcher Says She's Been Fired from Google', *MIT Technology Review*, 3 December 2020, <https://www.technologyreview.com/2020/12/03/1013065/google-ai-ethics-lead-timnit-gebru-fired/>. [题名中译: 一位顶尖AI伦理研究者称自己遭Google解雇。]
- . 'In 2020, Let's Stop AI Ethics-Washing and Actually Do Something', *MIT Technology Review*, 27 December 2019, <https://www.technologyreview.com/2019/12/27/57/ai-ethics-washing-time-to-act/>. [题名中译: 2020年, 让我们停止AI伦理洗白, 真正做点事情。]
- Hare, Julie. 'Wage Theft Is "Systemic": 21 Universities under Investigation', *Australian Financial Review*, 20 October 2021, <https://www.afr.com/work-and-careers/education/wage-theft-is-systemic-21-universities-under-investigation-20211020-p591kw>. [题名中译: 克扣工资是“系统性的”: 21所大学接受调查。]
- Hoffmann, Anna Lauren. 'Terms of Inclusion: Data, Discourse, Violence', *New Media & Society*, 16 September 2020, 146144482095872. [题名中译: 纳入的条件: 数据、话语与暴力。]
- hooks, bell. *Talking Back: Thinking Feminist, Thinking Black*, London: Routledge, 2014. [题名中译: 回嘴: 以女性主义思考, 以黑人立场思考。]
- Hu, Lily. 'Tech Ethics: Speaking Ethics to Power, or Power Speaking Ethics?' *Journal of Social Computing* 2, (3, 2021): 238–48. [题名中译: 科技伦理: 向权力讲伦理, 还是权力在讲伦理?]
- Johnson, Bobbie, and Gideon Lichfield. 'Hey Google, Sorry You Lost Your Ethics Council, So We Made One for You', *MIT Technology Review* (blog), 8 April 2019, <https://medium.com/mit-technology-review/hey-google-sorry-you-lost-your-ethics-council-so-we-made-one-for-you-28ee6c33576a>. [题名中译: 嘿, Google, 很遗憾你失去了伦理委员会, 所以我们给你组建了一个。]
- Manjoo, F. 'Silicon Valley Has an Arrogance Problem—WSJ', *Wall Street Journal*, 3 November 2013, <https://www.wsj.com/articles/SB10001424052702303661404579175712015473766>. [题名中译: 硅谷存在傲慢问题——《华尔街日报》。]
- McCarthy, Gregory Michael, and Jayasuriya, Kanishka. 'A Review into How University Research Works in Australia Has Just Begun—It Must Confront These 3 Issues', *The Conversation*, 20 September 2022, <http://theconversation.com/a-review-into-how>

- university-research-works-in-australia-has-just-begun-it-must-confront-these-3-issues-190551. [题名中译: 澳大利亚大学科研运作审查刚刚开始——它必须面对这三个问题。]
- Moreton-Robinson, Aileen. *Talkin' Up to the White Woman: Aboriginal Women and Feminism*, Brisbane: University of Queensland Press, 2000. [题名中译: 向白人女性直言: 原住民女性与女性主义。]
- Morozov, Evgeny. *To Save Everything, Click Here*, New York: PublicAffairs, 2014. [题名中译: 要拯救一切, 请点击这里。]
- Moten, Fred, and Stefano Harney. 'The University and the Undercommons', *Social Text* 22 (2, 2004): 101-15. [题名中译: 大学与底层公地。]
- Munn, Luke. 'The Uselessness of AI Ethics', *AI and Ethics*, 23 August 2022. [题名中译: AI 伦理的无用性。]
- Noble, Safiya. *Algorithms of Oppression: How Search Engines Reinforce Racism*, New York: NYU Press, 2018. [题名中译: 压迫的算法: 搜索引擎如何强化种族主义。]
- . 'Google Search: Hyper-Visibility as a Means of Rendering Black Women and Girls Invisible—InVisible Culture', *InVisible Culture*, no. 19 (October 2013), <https://ivc.lib.rochester.edu/google-search-hyper-visibility-as-a-means-of-rendering-black-women-and-girls-invisible/>. [题名中译: Google搜索: 过度可见性如何使黑人女性与女孩隐形——《不可见文化》。]
- . 'Missed Connections: What Search Engines Say About Women.' *Bitch magazine*, Spring 2012, <https://www.bitchmedia.org/issue/54>. [题名中译: 错失的联系: 搜索引擎如何谈论女性。]
- Ochigame, Rodrigo. 'The Invention of 'Ethical AI': How Big Tech Manipulates Academia to Avoid Regulation', *The Intercept* (blog), 20 December 2019, <https://theintercept.com/2019/12/20/mit-ethical-ai-artificial-intelligence/>. [题名中译: “伦理AI”的发明: 大型科技企业如何操纵学术界以逃避监管。]
- Phan, Thao, Jake Goldenfein, Monique Mann, and Declan Kuch. 'Economies of Virtue: The Circulation of "Ethics" in Big Tech', *Science as Culture*, 4 November 2021, 1-15. [题名中译: 美德经济: 大型科技企业中“伦理”的流通。]
- Powles, Julia, and Helen Nissenbaum. 'The Seductive Diversion of 'Solving' Bias in Artificial Intelligence', *OneZero* (blog), 7 December 2018, <https://onezero.medium.com/the-seductive-diversion-of-solving-bias-in-artificial-intelligence-890df5e5ef53>. [题名中译: “解决”人工智能偏见所造成的诱人转移。]
- Rosenblat, Alex, and Tamara Kneese. 'Algorithmic Accountability', 无日期, 6. [题名中译: 算法问责。]
- Sweeney, Latanya. 'Discrimination in Online Ad Delivery: Google Ads, Black Names and White Names, Racial Discrimination, and Click Advertising', *Queue* 11 (3, 2013): 10-29. [题名中译: 在线广告投放中的歧视: Google广告、黑人姓名与白人姓名、种族歧视以及点击广告。]
- Tarnoff, Ben. 'Postscript to J. Khadijah Abdurahman's "A Body of Work That Cannot Be Ignored"', *Logic magazine*, 25 December 2021, <https://logicmag.io/beacons/a-body->

of-work-that-cannot-be-ignored/. [题名中译: J. Khadijah Abdurahman 〈不可忽视的一系列工作〉的后记。]

Turner, Fred. *From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism* (插图版) Chicago: University of Chicago Press, 2008. [题名中译: 从反文化到赛博文化: Stewart Brand、全球网络与数字乌托邦主义的兴起。]

Upturn. ‘Civil Rights, Big Data and Our Algorithmic Future: A September 2014 Report on Social Justice and Technology’, 2014, <https://bigdata.fairness.io/wp-content/uploads/2015/04/2015-04-20-Civil-Rights-Big-Data-and-Our-Algorithmic-Future-v1.2.pdf>. [题名中译: 民权、大数据与我们的算法未来: 2014年9月关于社会正义与技术的报告。]

Vinsel, Lee. ‘You’ re Doing It Wrong: Notes on Criticism and Technology Hype’, *Medium* blog, 1 February 2021, <https://sts-news.medium.com/youre-doing-it-wrong-notes-on-criticism-and-technology-hype-18b08b4307e5>. [题名中译: 你做错了: 关于批评与技术炒作的笔记。]

Wagner, Ben. ‘Ethics As An Escape From Regulation. From “Ethics-Washing” to Ethics-Shopping?’ in *Ethics As An Escape From Regulation. From ‘Ethics-Washing’ To Ethics-Shopping?*, 84–9. Amsterdam: Amsterdam University Press, 2018. [题名中译: 伦理作为逃避监管的手段: 从“伦理洗白”到伦理选购?; 伦理作为逃避监管的手段: 从“伦理洗白”到伦理选购?]

Whittaker, Meredith. ‘The Steep Cost of Capture’, *Interactions* 28 (6, 2021): 50–5. [题名中译: 俘获的高昂代价。]

Wooldridge, Adrian. ‘The Coming Tech-Lash’, *The Economist*, 18 November 2013, <https://www.economist.com/news/2013/11/18/the-coming-tech-lash>. [题名中译: 即将到来的科技反弹。]

Young, Meg, Michael Katell, and P.M. Krafft. ‘Confronting Power and Corporate Capture at the FAccT Conference’ in *2022 ACM Conference on Fairness, Accountability, and Transparency*, 1375–86. Seoul Republic of Korea: ACM, 2022. [题名中译: 在FAccT会议上直面权力与企业俘获; 2022年ACM公平、问责与透明会议。]

原注

[1] Elettra Bietti, ‘From Ethics Washing to Ethics Bashing: A View on Tech Ethics from Within Moral Philosophy’, *SSRN Electronic Journal*, 2021: 211 & 217. [题名中译: 从伦理洗白到伦理抨击: 从道德哲学内部看科技伦理。] 返回

[2] Bietti, ‘From Ethics Washing to Ethics Bashing: A View on Tech Ethics from Within Moral Philosophy’. [题名中译: 从伦理洗白到伦理抨击: 从道德哲学内部看科技伦理。] 返回

[3] Thao Phan, Jake Goldenfein, Monique Mann, and Declan Kuch, ‘Economies of Virtue: The Circulation of “Ethics” in Big Tech’, *Science as Culture*, 4 November 2021, 1–15. [题名中译: 美德经济: 大型科技企业中“伦理”的流通。] 返回

- [4] Meredith Whittaker, ‘The Steep Cost of Capture’ , *Interactions* 28, no. (6) (2021): 50–5. [题名中译: 俘获的高昂代价。] 返回
- [5] 例如, 参见 Abeba Birhane, Elayne Ruane, Thomas Laurent, Matthew S. Brown, Johnathan Flowers, Anthony Ventresque, and Christopher L. Dancy, ‘The Forgotten Margins of AI Ethics’ in *2022 ACM Conference on Fairness, Accountability, and Transparency*, 948–58. FAccT’ 22. New York, NY, USA: Association for Computing Machinery, 2022; Ben Gansky, and Sean McDonald, ‘CounterFAccTual: How FAccT Undermines Its Organizing Principles’ in *2022 ACM Conference on Fairness, Accountability, and Transparency*, 1982–92. Seoul Republic of Korea: ACM, 2022; Young, Meg, Michael Katell, and P.M. Krafft. ‘Confronting Power and Corporate Capture at the FAccT Conference’ in *2022 ACM Conference on Fairness, Accountability, and Transparency*, 1375–86. Seoul Republic of Korea: ACM, 2022. 进一步讨论, 参见本书《退出者》。[题名中译: AI伦理被遗忘的边缘; CounterFAccTual: FAccT如何破坏自身组织原则; 在FAccT会议上直面权力与企业俘获; 2022年ACM公平、问责与透明会议。] 返回
- [6] Bietti, ‘From Ethics Washing to Ethics Bashing’ , 211 & 217. [题名中译: 从伦理洗白到伦理抨击。] 返回
- [7] Ben Wagner, ‘Ethics as an Escape from Regulation. From “Ethics-Washing” to Ethics-Shopping?’ in *Being Profiled*, Amsterdam: Amsterdam University Press, 2018, 84–9. [题名中译: 伦理作为逃避监管的手段: 从“伦理洗白”到伦理选购? ; 被画像。] 返回
- [8] Bietti, ‘From Ethics Washing to Ethics Bashing’ . [题名中译: 从伦理洗白到伦理抨击。] 返回
- [9] Rashida Richardson的观点, 见Bobbie Johnson and Gideon Lichfield, ‘Hey Google, Sorry You Lost Your Ethics Council, So We Made One for You’ , *MIT Technology Review* blog, 8 April 2019, <https://medium.com/mit-technology-review/hey-google-sorry-you-lost-your-ethics-council-so-we-made-one-for-you-28ee6c33576a>. [题名中译: 嘿, Google, 很遗憾你失去了伦理委员会, 所以我们给你组建了一个。] 返回
- [10] Jake Metcal的观点, 见Johnson & Lichfield, ‘Hey Google, Sorry You Lost Your Ethics Council, So We Made One for You’ . [题名中译: 嘿, Google, 很遗憾你失去了伦理委员会, 所以我们给你组建了一个。] 返回
- [11] Adam Greenfield的观点, 见Johnson & Lichfield, ‘Hey Google, Sorry You Lost Your Ethics Council, So We Made One for You’ . [题名中译: 嘿, Google, 很遗憾你失去了伦理委员会, 所以我们给你组建了一个。] 返回
- [12] Julia Powles and Helen Nissenbaum, ‘The Seductive Diversion of ‘Solving’ Bias in Artificial Intelligence’ , *OneZero* blog, 7 December 2018, <https://onezero.medium.com/the-seductive-diversion-of-solving-bias-in-artificial-intelligence-890df5e5ef53>. [题名中译: “解决”人工智能偏见所造成的诱人转移。] 返回
- [13] Lily Hu, ‘Tech Ethics: Speaking Ethics to Power, or Power Speaking Ethics?’ *Journal of Social Computing*, 2 (3, 2021): 240. [题名中译: 科技伦理: 向权力讲伦理, 还是权力在讲伦理?] 返回

- [14] Luke Munn, 'The Uselessness of AI Ethics', *AI and Ethics*, 23 August 2022: 无页码〔题名中译：AI伦理的无用性。〕 返回
- [15] Karen Hao, 'In 2020, Let's Stop AI Ethics-Washing and Actually Do Something,' *MIT Technology Review*, <https://www.technologyreview.com/2019/12/27/57/ai-ethics-washing-time-to-act/>.〔题名中译：2020年，让我们停止AI伦理洗白，真正做点事情。〕 返回
- [16] 参见Phan et al, 'Economies of Virtue: The Circulation of "Ethics" in Big Tech' and Whittakker, 'The Steep Cost of Capture' .〔题名中译：美德经济：大型科技企业中“伦理”的流通；俘获的高昂代价。〕 返回
- [17] 澳大利亚研究理事会的DECRA（Discovery Early Career Researcher Award，发现项目青年研究者奖），是澳大利亚职业生涯早期学者能够申请的最具声望的资助，平均成功率为17%。2022年略有改善，成功率为19.7%（参见Australian Research Council, "Selection Report: Discovery Early Career Research Award 2022 | Australian Research Council"（《遴选报告：2022年发现项目青年研究者奖 | 澳大利亚研究理事会》），Australian Government | Australian Research Council, 2022, <https://www.arc.gov.au/funding-research/funding-outcome/selection-outcome-reports/selection-report-discovery-early-career-research-award-2022>）。有关澳大利亚研究理事会与资助的进一步讨论，参见本书Richardson所撰章节。 返回
- [18] Gregory Michael McCarthy and Kanishka Jayasuriya, 'A Review into How University Research Works in Australia Has Just Begun—It Must Confront These 3 Issues', *The Conversation*, 20 September 2022, <https://theconversation.com/a-review-into-how-university-research-works-in-australia-has-just-begun-it-must-confront-these-3-issues-190551>.〔题名中译：澳大利亚大学科研运作审查刚刚开始——它必须面对这3个问题。〕 返回
- [19] Julie Hare, 'Wage Theft Is "Systemic": 21 Universities under Investigation', *Australian Financial Review*, 20 October, 2021, <https://www.afr.com/work-and-careers/education/wage-theft-is-systemic-21-universities-under-investigation-20211020-p591kw>.〔题名中译：克扣工资是“系统性的”：21所大学接受调查。〕 返回
- [20] Hu, 'Tech Ethics: Speaking Ethics to Power, or Power Speaking Ethics?', 240.〔题名中译：科技伦理：向权力讲伦理，还是权力在讲伦理？〕 返回
- [21] Hu, 'Tech Ethics: Speaking Ethics to Power, or Power Speaking Ethics?', 243.〔题名中译：科技伦理：向权力讲伦理，还是权力在讲伦理？〕 返回
- [22] Lee Vinsel, 'You're Doing It Wrong: Notes on Criticism and Technology Hype,' *Medium* blog, 1 February, 2021, <https://sts-news.medium.com/youre-doing-it-wrong-notes-on-criticism-and-technology-hype-18b08b4307e5>.〔题名中译：你做错了：关于批评与技术炒作的笔记。〕 返回
- [23] socialsciences.org.au/publications/annual-report-2021/. 返回
- [24] Joel Barnes. 'Defunding Arts Degrees Is the Latest Battle in a 40-Year Culture War', *The Conversation*, 3 July 2020, <http://theconversation.com/defunding-arts-degrees-is-the-latest-battle-in-a-40-year-culture-war-141689>.〔题名中译：撤销文科学位资助是一场持续40年文化战争的最新战役。〕 返回

- [25] Fred Moten and Stefano Harney, 'The University and the Undercommons', *Social Text* 22, (2, 2004): 101–15. [题名中译: 大学与底层公地。] 返回
- [26] Ben Tarnoff, 'Postscript to J. Khadijah Abdurahman' s "A Body of Work That Cannot Be Ignored" ', *Logic* magazine, 2021, <https://logicmag.io/beacons/a-body-of-work-that-cannot-be-ignored/>. [题名中译: J. Khadijah Abdurahman 〈不可忽视的一系列工作〉的后记。] 返回
- [27] Moten and Harney, 'The University and the Undercommons', 106. [题名中译: 大学与底层公地。] 返回
- [28] Alan Wooldridge, 'The Coming tech-lash', *The Economist*, 18 November 2013, <https://www.economist.com/news/2013/11/18/the-coming-tech-lash>. [题名中译: 即将到来的科技反弹。] 返回
- [29] 一个很好的例子是Peter Thiel出于自由意志主义兴趣而关注海上定居。一般性讨论参见F. Manjoo, "Silicon Valley has an Arrogance Problem" (《硅谷存在傲慢问题》), *Wall Street Journal*, 3 November 2013. 返回
- [30] 这是Google自1998年成立以来自述的著名使命。有关这一口号的进一步讨论, 参见本书《公开的秘密》。返回
- [31] Fred Turner, *From Counterculture to Cyberculture*, University of Chicago Press: Chicago, 2006; Evgeny Morozov, *To Save Everything, Click Here*, Public Affairs: New York City, 2013b. [题名中译: 从反文化到赛博文化; 要拯救一切, 请点击这里。] 返回
- [32] John Burn-Murdoch, 'The problem with algorithms: magnifying misbehaviour', *Guardian*, 14 August 2013. [题名中译: 算法的问题: 放大不当行为。] 返回
- [33] Luke Dormehl, 'Algorithms Are Great and All, But They Can Also Ruin Lives', *Wired*, 19 November 2014, <https://www.wired.com/2014/11/algorithms-great-can-also-ruin-lives/>. [题名中译: 算法固然出色, 但也可能毁掉人生。] 返回
- [34] Cynthia Dwork, M. Hardt, T. Pitassi, O. Reingold, & R. Zemel, 'Fairness Through Awareness', 2011, <https://arxiv.org/abs/1104.3913>. [题名中译: 通过觉察实现公平。] 返回
- [35] Latanya Sweeney, 'Discrimination in Online Ad Delivery: Google Ads, Black Names and White Names, Racial Discrimination, and Click Advertising', *Queue* 11 (3, 2013): 10–29. 8. [题名中译: 在线广告投放中的歧视: Google广告、黑人姓名与白人姓名、种族歧视以及点击广告。] 返回
- [36] Kate Crawford, 'The Hidden Biases in Big Data', *Harvard Business Review*, 1 April 2013, <https://hbr.org/2013/04/the-hidden-biases-in-big-data>. [题名中译: 大数据中的隐蔽偏见。] 返回
- [37] Nicolas Diakopoulos, 'Race Against the Algorithms', *The Atlantic*, 4 October 2013, <https://www.theatlantic.com/technology/archive/2013/10/rageagainst-the-algorithms/280255/>. [题名中译: 与算法赛跑〔原注题名如此〕。] 返回
- [38] Safiya Noble, 'Missed Connections: What Search Engines Say About Women', *Bitch* magazine, Spring 2012 (54), <https://www.bitchmedia.org/issue/54>; S. Noble, 'Google

- Search: Hyper-visibility as a Means of Rendering Black Women and Girls Invisible—InVisible Culture’, *Invisible Culture*, 19. <https://ivc.lib.rochester.edu/googlesearch-hyper-visibility-as-a-means-of-rendering-black-women-and-girls-invisible/>; Safiya Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism*, NYU Press: New York City, 2013. [题名中译: 错失的联系: 搜索引擎如何谈论女性; Google搜索: 过度可见性如何使黑人女性与女孩隐形——《不可见文化》; 压迫的算法: 搜索引擎如何强化种族主义。] 返回
- [39] Solon Barocas, and Andrew Selbst, ‘Big Data’ s Disparate Impact’, *California Law Review*, 104, (3), (June 2016): 671 - 732. [题名中译: 大数据的差别影响。] 返回
- [40] <https://civilrights.org/2014/02/27/civil-rights-principles-era-big-data>. 返回
- [41] Upturn, *Civil Rights, Big Data and Our Algorithmic Future: A September 2014 Report on Social Justice and Technology*, 2014, <https://bigdata.fairness.io/wp-content/uploads/2015/04/2015-04-20-Civil-Rights-Big-Data-and-Our-Algorithmic-Future-v1.2.pdf>. [题名中译: 民权、大数据与我们的算法未来: 2014年9月关于社会正义与技术的报告。] 返回
- [42] 参见 <https://governingalgorithms.org/>. 返回
- [43] Alex Rosenblat, Tamara Kneese, and danah boyd, ‘Algorithmic Accountability: A Workshop Primer Produced For: The Social, Cultural & Ethical Dimensions of “Big Data”’, 17 March 2014, New York, NY, <https://www.datasociety.net/pubs/2014-0317/AlgorithmicAccountabilityPrimer.pdf>. [题名中译: 算法问责: 为“大数据”的社会、文化与伦理维度工作坊编写的导论。] 返回
- [44] 例如, 参见 <https://www.fatml.org/>. 返回
- [45] 进一步讨论, 参见 Phan et al, ‘Economies of Virtue: The Circulation of “Ethics” in Big Tech’. [题名中译: 美德经济: 大型科技企业中“伦理”的流通。] 返回
- [46] Lily Hu, ‘Tech Ethics: Speaking Ethics to Power, or Power Speaking Ethics?’ *Journal of Social Computing* 2 (3, 2021): 243. [题名中译: 科技伦理: 向权力讲伦理, 还是权力在讲伦理?] 返回
- [47] J. Khadijah Abdurahman, ‘On the Moral Collapse of AI Ethics’ *Medium* blog, 8 December 2020, <https://upfromthecracks.medium.com/on-the-moral-collapse-of-ai-ethics-791cbc7df872>. Karen Hao, ‘A Leading AI Ethics Researcher Says She’ s Been Fired from Google’, *MIT Technology Review*, 3 December 2020, <https://www.technologyreview.com/2020/12/03/1013065/google-ai-ethics-lead-timnit-gebru-fired/>. [题名中译: 论AI伦理的道德崩塌; 一位顶尖AI伦理研究者称自己遭Google解雇。] 返回
- [48] Whittaker, ‘The Steep Cost of Capture’. [题名中译: 俘获的高昂代价。] 返回
- [49] 例如, 参见 Sara Ahmed, *On Being Included: Racism and Diversity in Institutional Life*, Duke University Press, 2012; bell hooks, *Talking Back: Thinking Feminist, Thinking Black*. 2 edition. Routledge, 2014; Aileen Moreton-Robinson, *Talkin’ Up to the White Woman: Aboriginal Women and Feminism*. Univ. of Queensland Press, 2000. [题名中译: 论被纳入: 制度生活中的种族主义与多样性; 回嘴: 以女性主义思考, 以黑人立场思考; 向白人女性直言: 原住民女性与女性主义。] 返回

- [50] 参见Sara Ahmed, “The Language of Diversity.” *Ethnic and Racial Studies* 30, no. 2 (March 2007): 235–56; Ruha Benjamin, *Race After Technology: Abolitionist Tools for the New Jim Code*. 1 edition. Cambridge, UK; Medford, MA: Polity, 2019; Carl Elliott, *White Coat, Black Hat: Adventures on the Dark Side of Medicine*. 1st edition. Beacon Press, 2010; Anna Lauren Hoffmann, “Terms of Inclusion: Data, Discourse, Violence.” *New Media & Society*, September 16, 2020, doi.146144482095872. [题名中译：多样性的语言；技术之后的种族：废除新吉姆代码的工具；白大褂，黑帽子：医学阴暗面的历险；纳入的条件：数据、话语与暴力。] [返回](#)
- [51] 参见Powles and Nissenbaum, ‘The Seductive Diversion of ‘Solving’ Bias in Artificial Intelligence’ ; Ochigame, ‘The Invention of “Ethical AI” ’ ; Gebru, ‘For Truly Ethical AI, Its Research Must Be Independent from Big Tech’ ; Abdurahman, ‘Fired for Speaking Out Against Anti-Blackness and Ethnic Cleansing: Open Letter to Cornell Tech’ and ‘On the Moral Collapse of AI Ethics’ ; Whittaker, ‘The Steep Cost of Capture’ ; Young, et al, ‘Confronting Power and Corporate Capture at the FAccT Conference’ . [题名中译：“解决”人工智能偏见所造成的诱人转移；“伦理AI”的发明；要有真正合乎伦理的AI，其研究必须独立于大型科技企业；因公开反对反黑人主义与族群清洗而被解雇：致康奈尔科技学院的公开信；论AI伦理的道德崩塌；俘获的高昂代价；在FAccT会议上直面权力与企业俘获。] [返回](#)
- [52] Pierre Bourdieu, *Science of Science and Reflexivity*, University of Chicago Press: Chicago, 2004. [题名中译：科学之科学与反思性。] [返回](#)
- [53] Jake Goldenfein and Daniel Griffin, ‘Google Scholar: Platforming the Scholarly Economy’ , *Internet Policy Review* 11 (3, 2022) (在线) . [题名中译：Google Scholar：学术经济的平台化。] [返回](#)

第一部分：主体

一分钱买你的思想？资本、共谋与AI伦理

Corinne Cath和Os Keyes

导言

随着人们对算法技术伤害的关注增加，以及旨在处理这些伤害的改良主义努力展开，一批主张更激进干预的研究与研究者也构成了批判性的侧翼——而且理由充分。人们对“伦理洗白”“伦理选购”和“伦理剧场”的忧虑，^[1]一再得到实例印证：一些公司试图淡化那些审视根植于AI商业模式和企业权力巩固过程之内在伤害的研究。^[2]这些令人担忧的动态，发生在科技行业游说力量不断增长的背景下，^[3]同时也伴随着新自由主义化的加深：政府与独立来源对学术研究的资助减少，并被产业资金取代。这种直接希望得到有利于商业之结果的企业资助，其日益重要的作用在AI伦理学术研究中尤为明显。^[4]

这种令人担忧的动态，提出了一系列关于研究AI技术伦理究竟意味着什么问题。它使我们追问，谁“拥有”AI伦理，^[5]并由此追问：谁有责任维持学术研究中的伦理标准？同样，私营产业在科研资助中的作用，也引出了企业如何以微妙而有害的方式向学术工作施压的问题，即便这些资金在形式上被称为“不附带任何条件”。最后同样重要的是，学术资助格局的变化也带来了新的挑战：产业影响如何作用于下一代学者——例如本文作者——的研究议程与职业轨迹。

我们提出的核心问题是，一个人在哪里、又如何划线——这一问题既是描述性的，也是规范性的。本章借鉴自我民族志方法，^[6]并以故事形式呈现我们的经历，考察这两个方面；两位作者各自以不同方式卷入产业资助这一现实，并对自己的经验作出反思。我们考察自己的工作如何受到产业资助塑造，如何针对自己的思想在何时、以何种方式被“一分钱买走”协商个人底线，以及这些协商与界线如何随时间演变。通过个体反思与相互对话，我们问自己，也问彼此：

在研究依赖产业资助的学术环境中开展AI工作，其民族志现实能够带来哪些教训？

回答这个众多研究者都要面对的问题，将有助于我们提出新的见解，说明企业权力如何在AI伦理学术研究的情境中发挥作用。

我们认为，尽管人们可以、也经常会在“常规”与“批判性”的数据伦理视角及行动者之间划定边界，但现实往往更复杂。参与并不意味着毫无疑问，Su、Lazar和Irani已经对此作了精到记录。^[7]同样，批评也不能避免共谋。^[8]在每一种视角与立场背后，都是人：他们在资源有限、企业俘获的研究环境中，摸索自己与自身价值观的关系，以及与这些价值观的疏离。^[9]

即使没有这样的环境，只要必须与持有不同利益和观点的他人及组织共同工作、建立关系，某种程度的妥协也就不可避免。正如Nick Seaver所言，伦理的一个重要场所，正是“人们如何在看似相互竞争的价值之间协商”。^[10]本章阐明这种协商在实践中可能是什么样子，以及对于那些试图在学术环境和产业资助的AI研究中，理解如何开展伦理工作、又如何与伦理发生关系的人来说，可以从中学到什么。

在整个反思过程中，两个主题不断浮现。其一，是企业间接塑造研究的权力：它们并不设定研究内容，而是划出可接受批评的边界。其二，是产业资助使哪些关系成为可能，又使哪些关系无法建立。我们两人都是博士研究者，开始研究的时间相隔三年、地点相距4,488英里，却在共同经验中发现这两个不同主题，以及女性主义拒绝的作用。^[11]这说明企业权力的影响何其广泛、对研究有何含义，也指向下一代通过产业资金开展AI伦理研究的学者所拥有的可能性。

本章安排如下。首先，我们聚焦于AI伦理中的资本、共谋与衔接工作，将自身工作和经历置于近期关于如何在学术界处理价值问题的文献之中。由此，我们展示自己作为初级学者从事AI伦理研究的个体经验，与文献所记录的研究遭到俘获、以及通过研究进行抵抗的更广泛动态之间的联系。我们表明，许多处于AI与伦理交汇点上的学者，其工作受到当代学术界资助现实的塑造；而试图简单逃离这些现实的努力，本身也纠缠于权力问题之中。其次，我们依据自身研究经历，介绍两个简短的自我民族志故事，采用的是Sotiropoulou和Cranston所使用的意义。^[12]

Corinne考察产业资助对于AI伦理研究的微妙权力动态，以及对下一代学者的影响。Os则关注资助如何塑造研究者建立的关系网络，以及由此仍然无法触及的关系。随后，我们让各自经验相互对话，进一步讨论AI伦理如何关联权力问题——这是本书的核心主题。最后我们认为，鉴于学术界与产业利益及网络的深度纠缠，它越来越难以成为对AI进行批判性思考的堡垒。如果学术界曾经真能发挥这种作用，^[13]那么上述结论便要求我们彻底重新思考AI伦理研究的资助方式，并在学术界内外建立新的反制权力与关怀网络。

定位我们：围绕AI伦理的持续辩论

描述和／或划定AI伦理边界的努力，往往依赖“浪潮”这一观念，即先后出现、各不相同的学术与思想路线。^[14]这种做法虽然常见，但我们担心它会使那些“不合时宜”的学术工作变得不可见、不可理解：它们也许太激进、出现太早，也许太常规、来得太迟。^[15]我们更愿意按照批评形式区分AI伦理研究：一类是内部批判，试图“从[学术]自身内部”改变学术；另一类是内在批判，虽参与被分析的实践，却依靠“超越情境的主张”。^[16]

内部批判在现有治理形式的规范内运作。在AI伦理中，它体现为自愿或非自愿伦理守则的提案，^[17]体现为将“算法公平”概念付诸实践并引导AI研究者朝此努力的

提案，也体现为尝试制定具有约束力的法规——这些法规隐含地进一步推进了许多政治人物、学者和商业领袖所信奉的AI不可避免论。^[18]这种批评往往受到企业资助者青睐，因此在AI伦理讨论中占据了大量空间。

与此相反，内在批判借鉴更广泛的规范性理解，希望带来的不仅是改良效果，而且是变革效果。在AI领域，这常被称为“批判性数据研究”：它不仅质疑算法系统的结果，也质疑系统设计所依赖的前提、使系统成为可能的政治经济条件，以及AI同时依赖并促成的话语文化。^[19]例如，Green和Vijoen概述了算法思维及其技术形式主义的局限，指出AI造成的许多伤害发生在这些边界之外，需要从社会层面、而非技术层面加以处理。^[20]同样，在近期为欧洲数字权利组织EDRI撰写的一份报告中，^[21]作者认为，必须处理AI权力失衡的根源。Powles和Nissenbaum也指出，通过专注于从内部改进技术而服从于AI，不仅限制了关于技术作用的批判性对话，也限制了伦理、法律与媒体在遏制AI伤害方面的作用。^[22] Powles、Nissenbaum、Balayn和Gürses等批判性学者认为，聚焦于“对系统进行技术性去偏见”，虽然流行，却不充分，而且把AI的社会技术影响过度简化成了技术设计的函数。他们强烈指出，这些内部批判和改进之所以诱人，反映的是其简单易行，而非它们适合处理AI不可避免地引出的权力与治理问题。（译注：原文此处将Viljoen误拼为Vijoen。）

可以理解的是，这些关注AI社会技术性质的研究者，也把目光转向了“AI伦理本身”。他们认为，常规AI伦理受到AI产业领导者的经济、社会和政治利益束缚，并且与这些利益存在根本联系。由于这种内在批判聚焦于解构支撑AI系统炒作的权力关系，它较不可能获得产业资助。

然而，人们日益需要这种凸显AI伦理“做了什么”的内在批判。这种批判可以成为AI伦理领域的元分析，并反对那种狭隘的批评：后者回避了对支撑自身的政治经济现实作结构性改变。例如，研究者担心，“公平”或伦理守则的提案可能构成“伦理洗白”，^[23]为有害的组织与过程披上一层合乎伦理的幻象。同样，人们也十分担忧：常规AI伦理观念以及从事实践的人，正在促成不公正未来的过程中共谋。科技公司与遭其伤害的社群之间，在政治经济权力上的巨大不平等，导致该领域即便还没有受到“监管俘获”，也已处在这种风险之中。^[24]

在这道往往充满紧张的分界线两侧，存在的不只是原则与实践，还有人。与古典社会学的看法相反，这些人并非对自己参与并推进的广泛过程毫无认识的“文化傻瓜”。相反，他们是具有能动性的存在，常常意识到这种来回拉扯，并且——正如Su等人与Cath分别展示的——^[25]会对它作出回应。于是，问题转变为：这些回应可能是什么样子；人们已经在如何处理这些张力，以及如何以“更好”的方式处理。

我们两位作者对此再清楚不过，因为我们都是AI伦理领域的青年学者。Corinne Cath最近在英国完成了博士学位，Os Keyes目前是在美国攻读博士学位的候选人。攻读博士之前，Corinne曾在政治领域及多个非营利组织工作；Os则曾在营利和非营利组织担任数据科学家。我们的研究主题分别是公民社会如何参与技术治理辩论，以及支配结构，它们反映了各自的个人背景、兴趣与身份。我们承认自身作为酷儿、跨性别、白人和欧洲学者的位置性，也认识到这些背景与我们的研究、与我们寻求创造AI伦理知识的方式之间，有着无法割裂的联系。本文始终借鉴自我民族志的认识论与方法，说明有关AI伦理的知识无法与从事研究的人分开，也无法与其专业知识得到何种认识论或物质支持分开。以下自我民族志故事，是将我们的经历与对话综合写成的叙述，大体呈现我们遇到的问题类型。^[26] 这些汇集的经验，展现产业对AI伦理研究的影响如何广泛渗透于我们博士研究所处的不同背景与时期。

Corinne的故事：“不附带条件”的产业研究资助何以不可能

2016年，我在欧洲一所以古老著称的大学开始攻读博士。在博士学习期间，我修读并参与讲授了校内不同教授的课程，他们也让我加入正在进行的研究项目。能够参与他们的研究，既令人兴奋，也让我感到受到认可。由于我的大学靠近伦敦科技圈，加上我研究计算文化的人类学背景，我参与的一些工作，自然包括为知名大型社交媒体公司开展研究。这些工作往往围绕特别棘手的主题，例如大规模使用AI进行内容审核的伦理。有时，这些研究会经过同行评审并面向公众；另一些时候，则供公司内部使用。有时，与我合作的资深研究者会获得这些科技公司的经费开展研究。

我个人有兴趣参与这种由产业推动的工作；它让我得以窥见机器内部，了解大型科技公司内部决策由哪些优先事项和政治因素驱动。研究的聚焦性质，以及与我合作的资深研究者和公司之间的信任关系，意味着分享的信息更详细，我们交谈的对象也较少保留。这些经历对我这样研究计算社群与文化的人类学者而言，提供了重要见解。我当时认为，即使无法明确写下自己所见所闻，做这类工作也会加深我对科技行业的整体理解，从而磨砺批评。我如今仍这样认为，但回头看，我也看到这类研究在提高研究者分析伤害能力的同时，如何扩充了这些公司反驳批评者的工具箱。

在我参与的产业项目中，明显缺席的是对科技行业某些最为人熟知的侵蚀性和有害影响的直接探讨：它以监控为基础的数据收集与商业模式。考虑到学术界已经对这种商业模式作出批评，^[27] 而且它又是许多后续伦理关切的根源，这种缺席令人惊讶。当我提出涉及这些商业模式的建议时，资深研究者往往成为阻止此类批评进入研究的第一道防线。我得到的回应总是类似这样：“我们不能告诉这家公司要改变

商业模式，虽然我们都知道它是问题的一部分，因为他们不会放弃自己的摇钱树。”

我曾多次询问，我们与公司之间的研究资助条款中，是否有什么规定使我们不能突出这些结构性问题。我得到许多回答，归根结底都是下面这段话的变体：“没有，我们想说什么就能说什么。这笔资金不附带任何条件。但我们不会告诉他们改变商业模式。”

“不附带任何条件”这个说法，意指资助大概不会对研究内容提出要求。只有在对这句话作最狭隘解释时，这种说法才成立。学者与产业资助者签署的协议，或许没有明文限制批评的性质，也没有限制将批评指向支撑产业成功的结构性力量。然而，围绕研究范围的默契——包括开展研究时不转向内在批判的惯例——确实构成了真实限制。在研究开始之前的许多环节，产业资助者所作的决定，意味着他们更可能邀请愿意改善产业现状、而非“颠覆”它的学者。

对于那些希望通过与合作，提出内部批判而非内在批判的人，学术界当然有其位置。我并不批评这一选择。然而，问题在于通过资助施加的无声权力，因为公司并不会平等地向学者分配经费。相反，我从参与这类工作的经历中认识到，不同公司会寻找研究议程与自身企业激励一致的学者。换言之，它们选择那些留在内部批判边界以内的研究。当然，这条规律也有例外，我并不是说，所有接受产业资助的学者都被阻止作出尖锐的结构性批评。尽管如此，我亲眼见过科技公司如何采取策略性放大和抬高地位的手段。Whittaker将这种动态描述为：“产业抬高自己最弱的批评者。”^[28]

产业能够、也确实通过建立并强化研究者网络来塑造研究：这些研究者提出温和、稳定、可预测的批评，不会过度聚焦于颠覆现状。产业对于研究议程的大部分权力，就存在于它与某些学者建立的关系之中。这种动态对初级学者有直接影响。产业资助塑造研究，并非直接命令产生某些结果，而是放大那些产业愿意回答的问题。它由此塑造何种批评算合理的边界，以及产业应当、或不必为哪些伤害负责。同时，资助这种内部批评，让公司看起来认真参与了学术研究，却也使它们能够把那些被视为过于激进、过于批判或过于大胆的学者排挤到一旁。

近期有多个超出我直接经历之外的例子，显示了这种动态，并支持我的分析。例如，Google一方面资助一个大学智库，后者积极批评对大型科技企业实施反垄断监管；^[29]另一方面，它又解雇内部与外部的学术批评者。^[30]我们每个人都需要判断，自己对这种现实能接受到什么程度，并让接受产业资助的决定扎根于一套有力、明确的政治立场之中，以便仔细衡量共谋或拒绝的影响——不仅对我们自己，而且对它鼓励我们建立何种关系、与谁建立、为了谁建立关系的影响。Os将在下一节继续讨论这一关系建立问题，以及它如何体现于产业资助塑造研究议程的能力之中。

Os的故事：“我们所有的关系”，或拓展人脉如何扭曲关系网络

2019年初，我获得了微软Ada Lovelace奖学金。这项奖学金很有声望：约100名决赛候选人中——只是大约，因为我当时忙着克制紧张得想吐的感觉，顾不上数——最后有五人获奖，候选人来自几乎所有STEM领域。但即使不谈声望，这也意义重大。它提供三年的学费和工资，因此，也提供三年专注研究的时间。

与Corinne的经历不同，这项奖学金不伴随赞助方持续介入。它的期待只是：我做自己感兴趣的研究，并保持在读身份。但缺乏干预并不意味着，它唯一的后果就是让我得以研究。还有两个后果值得强调：奖学金对微软的影响，以及它对我在研究中的个人与职业关系的影响。

它对微软的影响很快就显现出来。我的专业研究领域之一是人脸识别（FRT），我坚信这项技术应该被用木桩刺穿心脏，再丢在十字路口。到了这个时候，厌恶人脸识别已经成了学术家族的传统：我的一位导师的导师十年前就为此担忧；^[31]另一位则在二十年前就担忧过。^[32]因此，当我来到华盛顿州议会，以专家证人身份说明监管人脸识别的必要性，却发现反对我们的是……微软时，这是一场令人不快的意外——尽管在我最犬儒的时刻，它也并非出乎意料。他们的论点有多个方面，但其中一个关键部分是：监管并不必要，因为可以相信他们自己会处理人脸识别的各种偏见。具体而言，人脸识别需要的是来自多个边缘化社群、具有多样性和专业知识的利益相关者参与。换言之：像我这样的人。

当然，没有哪家公司是铁板一块。我完全愿意相信，资助这项奖学金的微软研究院中，有相当一部分人反对人脸识别，尽管整家公司在倡导并销售它。而且，那位律师显然并不是特指我；对方根本不知道我是谁。但事实是，微软之所以能够花钱养一个想把其某项产品付之一炬的人，原因在于它觉得这样做对自己有用。这样，它就能假装解决方案在“内部”，就能收编批评。它让奖学金获得者处在Fred Moten猛烈抨击批判学术界时所指出的同一角色之中：被赋予完善技术科学的任务，并非替代常规做事方式，而是“使其完成的尝试”。^[33]这并不是说，获得资助的批判性工作不可能有效；而是说，批评只是它的一种效果。

即便这项奖学金没有以微小方式改变微软的公关策略，它也确实改变了我。实现这种改变的方式之一，是奖学金中一条常见的附加条款：除了津贴与学费，获奖者还可优先获得微软研究院的实习机会。表面上看，这是双赢：微软能够试图说服博士生，毕业后全职加入公司是多么令人愉快；而这些学生也终于能吃点方便面以外的东西。

但是，做这项工作、身处那个地方、与那些人相伴，也带来另一个后果。参加实习，意味着与人建立关系，意味着接触某些非常聪明、友善的研究者的成本大幅降低。我通过这些机会，结识了一些自己最喜欢的人，包括Corinne。但由于我们在

世的时间有限、能力也受限制，我们只能与有限的人建立关系。因此，这些机会在创造与某些人亲近的可能性时，也限制了与另一些人建立关系的可能性。奖学金让我接触到新的合作者，并扩大了在专业研究圈中互动的空间。但是，它也夺走并削弱了我与那些最初使我获得这项奖学金的人之间的联系，削弱了我与赋予研究动力的人和社群之间的联系。

讨论

通常，讨论部分把分析收拢起来，探究其底层主题，再将这些主题压缩为固定的教训、视角或实践启示。尝试撰写这一部分时，我们很难明确指向任何单一方向，原因很简单：这里没有简单答案。Jones指出，“选择……在于进行‘参与性的’还是‘共谋性的’研究”，^[34]但很难找到一个没有共谋的位置——如果这样的位置真的存在的话。

因此，我们没有照常写一个讨论部分，而是从Bellanova等人的对话方式中汲取灵感，^[35]也从自己写作过程的对话方式中汲取灵感。与Sotiropoulou和Cranston一样，^[36]我们认为，这种对话方式及其背后的友谊，正在建立学术界中的激进关怀伦理实践。我们需要这种实践，来批判性地反思、并在必要时抵抗产业资助对独立研究的影响。我们决定呈现一组不固定的视角：它们暧昧、不确定，有时相互紧张，就像我们所讨论的动态一样。

通过这些视角，我们对“在研究依赖产业资助的学术环境中开展AI工作，其民族志现实能够带来哪些教训？”这一问题，形成了两个回答。第一个主题，是由于产业资助的引导力量而被建立、以及从未被追求的网络与关系的重要性。第二个主题，是产业资助温和却有害的权力，以及它在个体和系统两个层面对批判性研究的侵蚀作用。

一、回答：直接说不

面对共谋困境，最本能的反应，是要求一个“不”：要求彻底拒绝共谋的纯洁性，要求研究者拒绝站在那个位置。这种观点的优势在于道德上的清晰，但它很快就变得复杂，正如Shotwell等研究者以及明确讨论女性主义拒绝概念的学者所阐明的。^[37]

但是，究竟谁在字面意义上负担得起这样做？在我们所担忧的大型营利企业之外，AI伦理研究越来越不稳定，而人们获得那些确实存在的资源的机会又并不相等。我们还从一些并不令人惊讶的零散经历中看到，说“是”的后果也因人而异。Corinne例子中的资深学者，能够承受与这类资助者合作带来的声誉成本；他们积累的权力，使人们更容易原谅其共谋，也使他们比初级学者更不容易承受后果。

其次，否则我们又能站在哪里？大学也有自身的道德暧昧，Phan等人早已指出这一点。^[38]转向社区团体和网络是理想选择，但这些团体往往缺少支持研究者的经费，而且还有许多更为迫切、关键的事情需要花钱。在这里，非产业资助者与政府资助者可以发挥关键作用，确保有许多由组织和个人构成、从事批判性工作的据点。近期一个很好的例子，是Timnit Gebru博士创立的分布式人工智能研究所（DAIR），由多个非营利基金会资助。我们需要更多这样的反制权力场所，也需要以一种承认将“非营利”直接等同于“有益于社群”所带来风险的方式，来构建这些场所。^[39]

二、回答：拿了钱就跑

Fred Moten说：“与大学之间唯一合乎伦理的关系，就是犯罪关系。”^[40]他的意思是，面对一个要求冷酷无情的工作伦理、以不公正前提运作，却只回馈少量资源的机构，研究者应当尽可能多地拿取。如果这句话适用于大学，那么它对科技行业的许多公司就更加适用。参与的张力在于，这些组织确实拥有我们想要的资源；拿了钱就跑，便解决了——或至少改变了——这种张力。

这种做法的问题在于，它是一种战术，而不是战略。你也许能拿到一两次产业资助就跑，但产业资助者最终会察觉。而且，在支持研究方面，即使激烈竞争的产业对手，也会互通消息。对于那些留意到同样五位教授不断被邀请加入当下时兴的伦理委员会、理事会或监督机制的人来说，这不该令人惊讶。拿钱就跑的名声最终会先于你本人传播，堵住可能的关系——无论是犯罪式的，还是其他关系。此外，拿着你的一分钱就跑，最终会导致一种情形：你被排除出可以得到资助的学者群体，实际上等于在说“不”；而由于声誉具有集体性质，你也可能被视为替自己的学生一并表态。

三、转换问题：女性主义拒绝

这两个答案之所以如此令人不满意，部分原因是根本不存在一种超越一切情境、令人满意的答案。但另一个原因在于，它们提出了*错误的问题*。两者问的都是：

“我，一个单独的人，独自能做什么？”以这种方式处理问题，就忽略了女性主义和／或美德实践的两个关键面向：它们往往是关系性的，涉及人的集合；而且，部分正因此，它们是实践。答案取决于具体情境，追求答案的过程会改变情境，转而要求我们重新评估正在进行的行动。

因此，我们最终落脚的地方，既不是把接受参与的代价作为独立答案，也不是把拒绝参与作为独立答案，而是Bonnie Honig所称的女性主义拒绝。^[41]实行女性主义拒绝，不是“说不”之后就完事。相反，它从承认我们从未处在权力关系之外开始。正因此，对于那些更深地嵌在这些权力结构中的人，我们负有关怀责任，要将

拒绝视为一种战术：有目的地拒绝，可能是暂时地拒绝，着眼于未来；或者，为了某种战术目的，根据具体情境继续保持纠缠，并与他人的选择相联系。

在研究者共谋于AI伦理的情形中，这会是什么样子？一种假设性的表现形式，是合作开展一项产业资助研究，同时明确理解公司愿意推动的改变有哪些边界，并把这些边界本身作为研究不可分割的组成部分。另一种，是在受到产业邀请时参与对话，亲自了解它们如何界定问题和解决方案，再利用这些知识，磨砺集体处理伤害的尝试。还有一种，是将直接受影响的社群纳入研究设计，推动科技公司所聆听声音的边界。或者，以教育者身份参与，利用合作勾勒现状之外的替代方案，争取通过内部研究促成改变。

研究设计有许多部分——从理论、数据、方法，到研究结果的可获取性——都应成为我们划定研究边界、接受产业资助时协商过程的一部分。因此，女性主义拒绝既不天然排斥拒绝资助，也不天然排斥接受资助；它只是要求把两者都视为战术，在意识到他人的选择、并与这些选择发生关系的情况下采取行动。此外，与Audra Simpson关于原住民拒绝的平行思路相联系，^[42]我们认为，这不仅要求我们作出选择，也要求持续审视这些选择。要以关系性的方式，^[43]观察我们的暂时答案在实践中有时间展开之后，带来了什么好处与伤害。这种反思需要在与同伴、学生和导师的可信关系之中、并通过这些关系发生，也就是通过“批判性学术友谊”发生。^[44]我们不应只是回答这个问题，而应重新回答它：从集体的视角界定集体需要什么，而不是“我独自能做什么”。

结语：我们的经历对研究现状及其未来说明了什么？

我们描绘的图景很复杂，包容着多重面向。然而，从我们作为青年学者研究AI技术伦理的经历中，仍可得出明确教训。作为定性研究者，我们说明了女性主义拒绝、友谊，以及从个体转向集体的重新定位，如何至少部分回答了有效抵抗的问题。我们还希望强调，对学术知识如何被创造进行反思，也是关键。问题的一部分，在于我们描述的许多动态都发生在暗处。对于产业关系，很难做到公开透明，这反过来又使我们难以确切理解它们如何影响研究。我们希望，通过展示产业影响渗透学术界的细微方式，能够进一步发展既有的描述：既描述产业的影响，也描述企业如何否认自己在钝化科技批评中的共谋。

我们的基层经验把企业俘获可能如何影响学术界独立思想具体化，从而加强了既有对其危险的批评。它也对批判性学术研究的未来提出了明确忧虑：很少有年轻学者能逃脱或避免加在其机构之上的结构性压力，即必须依产业的规则与机制行事。这种研究“新常态”具有深刻而广泛的后果，也提出了抵抗可以、又应当是什么样子的的问题。

关于有效抵抗，我们有一些想法。我们呼吁其他青年学者寻找志同道合的社群与友谊，以支持持续的批判性反思和对话，讨论个体决定对学术共同体的影响。同样，我们也请终身教职教师进一步示范并参与这种反思，因为他们既最接近实际运作，也最不容易因为坦诚公开而受到伤害。我们希望强调他人已经提出的呼吁：需要把科技劳动者和学者组织成能够抵抗企业压力的集体。^[45]

对于真正有志于变革、而不仅是改良的资助者，我们会说：看看现有网络之外。看看那些不符合典型学术机构、正式慈善组织和公司模式的项目与组织。看看那些不承诺大规模回报的项目。看看社区团体；看看小空间中的大改变。接受这样一个事实：你们的介入本身也可能受到必要的审视。此外，本文也向有意追究科技责任的政府发出行动呼吁，请它们认真考虑资助批判性学术研究。要找到好的研究，就需要关注哪些学者和问题目前被排除在产业资助范围之外，并找出那些从技术的社会影响与集体福祉出发提出研究问题的学者，而不是那些只关注如何让AI产业更合乎伦理这一狭窄问题、捡了芝麻丢了西瓜的学者。

资助披露

Corinne对本研究的贡献未获资助；Os获得微软Ada Lovelace奖学金资助。

参考文献

- Agre, Phil. 'Your Face is not a Bar Code: Arguments Against Automatic Face Recognition in Public Places', *Whole Earth* 106.2001 (2001): 74–77. [题名中译：你的脸不是条形码：反对在公共场所自动识别人脸的论证。]
- Algorithm Watch. 'AlgorithmWatch Forced to Shut down Instagram Monitoring Project after Threats from Facebook', *AlgorithmWatch* blog, 13 August 2021, <https://algorithmwatch.org/en/instagram-research-shut-down-by-facebook/>. [题名中译：AlgorithmWatch受到Facebook威胁后被迫关闭Instagram监测项目。]
- Balayn, Agathe, and Gürses, Seda. 'If AI Is the Problem, Is Debiasing the Solution?', Brussels Belgium: EDRI, 2021, <https://edri.org/our-work/if-ai-is-the-problem-is-debiasing-the-solution/>. [题名中译：如果AI是问题，去偏见是解决方案吗？]
- Bank, Max, Duffy, Felix, Leyendecker, Verena, and Silva, Margarida. 'The Lobby Network: Big Tech's Web of Influence In the EU', Corporate Europe Observatory and LobbyControl e.V, 2021, <https://corporateeurope.org/en/2021/08/lobby-network-big-techs-web-influence-eu>. [题名中译：游说网络：大型科技企业在欧盟的影响力之网。]
- Bellanova, Rocco, Kristina Irion, Katja Lindskov Jacobsen, Francesco Ragazzi, Rune Saugmann, and Lucy Suchman. 'Toward a Critique of Algorithmic Violence.' *International Political Sociology* (2021.)15 (1): 121–50. <https://doi.org/10.1093/ips/olab003>. [题名中译：走向对算法暴力的批判。]
- Birhane, Abeba. 'Algorithmic Injustice: A Relational Ethics Approach.' *Patterns* 2.2 (2021). [题名中译：算法不公正：一种关系伦理路径。]

- Cath, Corinne. ‘Governing Artificial Intelligence: Ethical, Legal and Technical Opportunities and Challenges’ , *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376 (2018). [题名中译: 治理人工智能: 伦理、法律与技术的机遇和挑战。]
- Cath, Corinne, and Jansen, Fieke. ‘Dutch Comfort: The Limits of AI Governance Through Municipal Registers’ , *arXiv e-prints* (2021): arXiv-2109. [题名中译: 荷兰式安慰: 通过市政登记册治理AI的局限。]
- Conquergood, Dwight. ‘Between Rigor and Relevance: Rethinking Applied Communication’ , *Applied communication in the 21st century* (1995): 79–96. [题名中译: 在严谨与相关性之间: 重新思考应用传播; 21世纪的应用传播。]
- Dencik, Lina. ‘The Datafied Welfare State: A Perspective from the UK’ , in Hepp, Andreas, Jarke, Juliane, and Kramp, Leif (eds), *The Ambivalences of Data Power: New Perspectives in Critical Data Studies*, UK: Palgrave Macmillan, 2022):145–66. [题名中译: 数据化福利国家: 来自英国的视角; 数据权力的矛盾性: 批判性数据研究的新视角。]
- Ellis, Carolyn, Adams, Tony E., and Bochner, Arthur P. ‘Autoethnography: An Overview’ , *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 12.1 (2010). [题名中译: 自我民族志概述。]
- Green, Ben, and Viljoen, Salomé. ‘Algorithmic Realism: Expanding the Boundaries of Algorithmic Thought’ in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, Barcelona, Spain: ACM, 2020: 19–31. [题名中译: 算法现实主义: 拓展算法思维的边界; 2020年公平、问责与透明会议论文集。]
- Hemmings, Clare. *Why Stories Matter*, Durham: Duke University Press, 2011. [题名中译: 故事为何重要。]
- Honig, Bonnie. *A Feminist Theory of Refusal*, Cambridge: Harvard University Press, 2021. [题名中译: 女性主义拒绝理论。]
- Incite! Women of Color Against Violence (eds). *The Revolution Will Not Be Funded: Beyond the Non-profit Industrial Complex*, Boston: South End Press, 2007. [题名中译: 革命不会得到资助: 超越非营利产业复合体。]
- Jaeggi, Rahel. *Critique of Forms of Life*, Cambridge: Harvard University Press, 2018. [题名中译: 生活形式批判。]
- Jobin, Anna, Ienca, Marcello, and Vayena, Effy. ‘The Global Landscape of AI Ethics Guidelines’ , *Nature Machine Intelligence* 1.9 (2019): 389–99. [题名中译: AI伦理指南的全球图景。]
- Jones, Richard G. ‘Putting Privilege Into Practice Through “Intersectional Reflexivity” : Ruminations, Interventions, and Possibilities’ , in *Reflections: Narratives of Professional Helping*, 2010. [题名中译: 通过“交叉性反思”将特权付诸实践: 思考、干预与可能性; 反思: 专业助人工作的叙事。]
- Lapowsky, Issie. ‘How Tech Giants Court and Crush the People Who Study Them’ , *Protocol—The People, Power and Politics of Tech* blog, 19 March 2021. <https://www.protocol.io/tech-giants-court-and-crush-the-people-who-study-them/>

- protocol.com/nyu-facebook-researchers-scraping. [题名中译：科技巨头如何拉拢并打压研究它们的人。]
- Metcalfe, Jacob, and Moss, Emanuel. 'Owning Ethics: Corporate Logics, Silicon Valley, and the Institutionalization of Ethics', *Social Research: An International Quarterly* 86.2 (2019): 449–76. [题名中译：占有伦理：企业逻辑、硅谷与伦理的制度化。]
- Metz, Cade and Wakabayashi, Daisuke. 'Google Researcher Says She Was Fired Over Paper Highlighting Bias in A.I.', *The New York Times*, 3 December 2020. [题名中译：Google研究者称因论文揭示AI偏见而被解雇。]
- Moten, Fred, and Harney, Stefan. 'The University and the Undercommons: Seven Theses', *Social Text* 22.2 (2004): 101–15. [题名中译：大学与底层公地：七条论纲。]
- O'Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, New York: Crown, 2016. [题名中译：数学毁灭性武器：大数据如何加剧不平等并威胁民主。]
- Pasquale, Frank. *The Black Box Society: The Secret Algorithms That Control Money and Information*, Cambridge: Harvard University Press, 2016. [题名中译：黑箱社会：控制金钱与信息的秘密算法。]
- . 'Machines Judging Humans: The Promise and Perils of Formalizing Evaluative Criteria.', in *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 2020. [题名中译：机器评判人类：评估标准形式化的希望与危险；AAAI/ACM人工智能、伦理与社会会议论文集。]
- Phan, Thao, et al. 'Economies of Virtue: The Circulation of 'Ethics' in Big Tech', *Science as Culture* 31.1 (2021): 1–15. [题名中译：美德经济：大型科技企业中“伦理”的流通。]
- Popiel, Pawel. 'The Tech Lobby: Tracing the Contours of New Media Elite Lobbying Power', *Communication Culture & Critique* 11.4 (2018): 566–85. [题名中译：科技游说集团：追踪新媒体精英游说权力的轮廓。]
- Reed-Danahay, Deborah. *Auto/Ethnography: Rewriting the Self and the Social*, Oxford; New York: Routledge, 1997. [题名中译：自我/民族志：重写自我与社会。]
- Rincón, Cami, Keyes, Os, and Cath, Corinne. 'Speaking from Experience: Trans/Non-Binary Requirements for Voice-Activated AI', in *Proceedings of the ACM on Human-Computer Interaction* 5.CSCW1 (2021): 132:1–132:27. [题名中译：从经验出发：跨性别与非二元性别者对语音激活AI的要求。]
- Seaver, Nick. 'CARE AND SCALE: Decorrelative Ethics in Algorithmic Recommendation', *Cultural Anthropology*. 36.3 (2021): 509–37. [题名中译：关怀与规模：算法推荐中的去相关伦理。]
- Shotwell, Alexis. *Against Purity: Living Ethically in Compromised Times*, Minneapolis: University of Minnesota Press, 2016. [题名中译：反对纯洁：在妥协时代合乎伦理地生活。]

- Simpson, Audra. ‘The Ruse of Consent and the Anatomy of “Refusal” : Cases from Indigenous North America and Australia’ , *Postcolonial Studies*, 20.1 (2017): 18–33. [题名中译: 同意的诡计与“拒绝”的剖析: 北美和澳大利亚原住民案例。]
- Sloane, Mona. ‘Inequality Is the Name of the Game: Thoughts on the Emerging Field of Technology, Ethics and Social Justice’ , Weizenbaum Conference, Berlin, Germany, 16–17 May 2019. [题名中译: 不平等才是游戏规则: 对技术、伦理与社会正义新兴领域的思考。]
- Sotiropoulou, Panagiota, and Cranston, Sophie. ‘Critical Friendship: An Alternative, ‘Care-full’ Way to Play the Academic Game’ , *Gender, Place & Culture* (2022): 1–22. [题名中译: 批判性友谊: 一种充满关怀的学术游戏替代玩法。]
- Star, Susan Leigh. ‘This is Not a Boundary Object: Reflections on the Origin of a Concept’ , *Science, Technology, & Human Values*, 35.5 (2010): 601–17. [题名中译: 这不是边界对象: 对一个概念起源的反思。]
- Su, Norman Makoto, Lazar, Amanda, and Irani, Lilly. ‘Critical Affects: Tech Work Emotions Amidst the Techlash’ in *Proceedings of the ACM on Human-Computer Interaction 5*. CSCW1 (2021): 1–27. [题名中译: 批判性情感: 科技反弹中的科技工作情绪。]
- Tarnoff, Ben. *The Making of the Tech Worker Movement*. New York: *Logic* magazine, 2019, <https://logicmag.io/the-making-of-the-tech-worker-movement/>. [题名中译: 科技劳动者运动的形成。]
- Taylor, Linnet. ‘Public Actors Without Public Values: Legitimacy, Domination and the Regulation of the Technology Sector’ , *Philosophy & Technology*, (2021):1–26. [题名中译: 没有公共价值的公共行动者: 正当性、支配与科技行业监管。]
- Wagner, Ben. ‘Ethics As An Escape From Regulation. From “Ethics-Washing” To Ethics-Shopping?’ in *Being Profiled*, Amsterdam University Press, 2018, 84–9. [题名中译: 伦理作为逃避监管的手段: 从“伦理洗白”到伦理选购? ; 被画像。]
- Wakabayashi, Daisuke. ‘Big Tech Funds a Think Tank Pushing for Fewer Rules. For Big Tech’ , *The New York Times*, 24 July 2020. [题名中译: 大型科技企业资助智库, 推动减少对大型科技企业的监管。]
- Whittaker, Meredith. ‘The Steep Cost of Capture’ , *Interactions*, 28.6 (2021): 50–5. [题名中译: 俘获的高昂代价。]

原注

- [1] 参见Mona Sloane, ‘Inequality is the Name of the Game: Thoughts on the Emerging Field of Technology, Ethics and Social Justice’ , Weizenbaum Conference, Berlin, 16–17 May 2019; 另见Ben Wagner, ‘Ethics As An Escape From Regulation. From “Ethics-Washing” To Ethics-Shopping?’ , in *Being Profiled*, Amsterdam: Amsterdam University Press, 2018; and Corinne Cath and Fieke Jansen, ‘Dutch Comfort: The Limits of AI Governance Through Municipal Registers.’ *arXiv e-prints* (2021): arXiv–2109. [题名中译: 不平等才是游戏规则: 对技术、伦理与社会正义新兴领域的思考; 伦理作为逃避监管的手段: 从“伦理洗白”到伦理选购? ; 被画像; 荷兰式安慰: 通过市政登记册治理AI的局限。] 返回

- [2] Algorithm Watch, ‘AlgorithmWatch Forced to Shut down Instagram Monitoring Project after Threats from Facebook’ . *AlgorithmWatch*, 13 August 2021, <https://algorithmwatch.org/en/instagram-research-shut-down-by-facebook/> and Issie Lapowsky, ‘How Tech Giants Court and Crush the People Who Study Them’ , Protocol—The People, Power and Politics of Tech, 19 March 2021, <https://www.protocol.com/nyu-facebook-researchers-scraping>. [题名中译: AlgorithmWatch受到Facebook威胁后被迫关闭Instagram监测项目; 科技巨头如何拉拢并打压研究它们的人。] 返回
- [3] Pawel Popiel, ‘The Tech Lobby: Tracing the Contours of New Media Elite Lobbying Power.’ *Communication Culture & Critique* 11.4 (2018): 566–85 and Max Bank, Felix Duffy, Verena Leyendecker, and Margarida Silva, ‘The Lobby Network: Big Tech’ s Web of Influence in the EU’, Corporate Europe Observatory and LobbyControl e.V, 2021, <https://corporateeurope.org/en/2021/08/lobby-network-big-techs-web-influence-eu>. [题名中译: 科技游说集团: 追踪新媒体精英游说权力的轮廓; 游说网络: 大型科技企业在欧盟的影响力之网。] 返回
- [4] Corinne Cath, ‘Governing Artificial Intelligence: Ethical, Legal and Technical Opportunities and Challenges’ , *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376 (2018) and Agathe Balayn and Seda Gürses, ‘If AI Is the Problem, Is Debiasing the Solution?’ Brussels, Belgium: EDRI, 2021, <https://edri.org/our-work/if-ai-is-the-problem-is-debiasing-the-solution/>. [题名中译: 治理人工智能: 伦理、法律与技术的机遇和挑战; 如果AI是问题, 去偏见是解决方案吗?] 返回
- [5] Jacob Metcalf and Emanuel Moss, ‘Owning Ethics: Corporate Logics, Silicon Valley, and the Institutionalization of Ethics’ , *Social Research: An International Quarterly* 86, (no. 2, 2019): 449–76. [题名中译: 占有伦理: 企业逻辑、硅谷与伦理的制度化。] 返回
- [6] Carolyn Ellis, Tony E. Adams, and Arthur P. Bochner, ‘Autoethnography: An Overview’ , *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 12.1 (2001) and Deborah Reed-Danahay, *Auto/Ethnography: Rewriting the Self and the Social*, Oxford; New York: Routledge, 1997. [题名中译: 自我民族志概述; 自我/民族志: 重写自我与社会。] 返回
- [7] Norman Makato Su, Amanda Lazar, and Lilly Irani, ‘Critical Affects: Tech Work Emotions Amidst the Techlash’ , *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021). [题名中译: 批判性情感: 科技反弹中的科技工作情绪。] 返回
- [8] Fred Moten and Stefano Harney, ‘The University and the Undercommons: Seven theses’ , *Social Text* 22.2 (2004). [题名中译: 大学与底层公地: 七条论纲。] 返回
- [9] Cath, ‘Governing Artificial Intelligence’ and Meredith Whittaker, ‘The Steep Cost of Capture’ , *Interactions* 28.6 (2021): 50–5. [题名中译: 治理人工智能; 俘获的高昂代价。] 返回
- [10] Nick Seaver ‘CARE AND SCALE: Decorrelative Ethics in Algorithmic Recommendation’ , *Cultural Anthropology* 36 (no. 3, 2021): 509–37. [题名中译: 关怀与规模: 算法推荐中的去相关伦理。] 返回

- [11] Bonnie Honig, *A Feminist Theory of Refusal*, Cambridge: Harvard University Press, 2021 and Audra Simpson, 'The Ruse of Consent and the Anatomy of "Refusal": Cases from Indigenous North America and Australia.' *Postcolonial Studies* 20.1 (2017): 18–33. [题名中译: 女性主义拒绝理论; 同意的诡计与“拒绝”的剖析: 北美和澳大利亚原住民案例。] 返回
- [12] Panagiota Sotiropoulou and Sophie Cranston, 'Critical Friendship: An Alternative, 'Care-ful' Way to Play the Academic Game.' *Gender, Place & Culture* (2022). [题名中译: 批判性友谊: 一种充满关怀的学术游戏替代玩法。] 返回
- [13] Thao Phan, et al., 'Economies of Virtue: The Circulation of "Ethics" in Big Tech', *Science as Culture* (2021): 121–35. [题名中译: 美德经济: 大型科技企业中“伦理”的流通。] 返回
- [14] Frank Pasquale, 'Machines Judging Humans: The Promise and Perils of Formalizing Evaluative Criteria', in *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, New York, 7–8 February 2020. [题名中译: 机器评判人类: 评估标准形式化的希望与危险; AAAI/ACM人工智能、伦理与社会会议论文集。] 返回
- [15] Clare Hemmings, *Why Stories Matter*, Durham: Duke University Press, 2011. [题名中译: 故事为何重要。] 返回
- [16] Rahel Jaeggi, *Critique of Forms of Life*, Cambridge: Harvard University Press, 2018. [题名中译: 生活形式批判。] 返回
- [17] 参见Anna Jobin, Marcello Ienca, and Effy Vayena, 'The Global Landscape of AI Ethics Guidelines', *Nature Machine Intelligence* 1.9 (2019). [题名中译: AI伦理指南的全球图景。] 返回
- [18] Cath & Jansen, 'Dutch Comfort'. [题名中译: 荷兰式安慰。] 返回
- [19] Lina Dencik, 'The Datafied Welfare State: A Perspective from the UK' in Andreas Hepp, Juliane Jarke, and Leif Kramp (eds) *The Ambivalences of Data Power: New Perspectives in Critical Data Studies*, Palgrave Macmillan UK: 2021; Cami Rincón, Os Keyes, and Corinne Cath, 'Speaking from Experience: Trans/Non-Binary Requirements for Voice-Activated AI', *Proceedings of the ACM on Human-Computer Interaction* (2021), and Linnet Taylor, 'Public Actors Without Public Values: Legitimacy, Domination and the Regulation of the Technology Sector', *Philosophy & Technology*, 2021. [题名中译: 数据化福利国家: 来自英国的视角; 数据权力的矛盾性: 批判性数据研究的新视角; 从经验出发: 跨性别与非二元性别者对语音激活AI的要求; 没有公共价值的公共行动者: 正当性、支配与科技行业监管。] 返回
- [20] Ben Green and Salomé Viljoen, 'Algorithmic Realism: Expanding the Boundaries of Algorithmic Thought' in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* 2020, Barcelona Spain, 2020: ACM. [题名中译: 算法现实主义: 拓展算法思维的边界; 2020年公平、问责与透明会议论文集。] 返回
- [21] Balayn and Gürses, 'If AI Is the Problem, Is Debiasing the Solution?'. [题名中译: 如果AI是问题, 去偏见是解决方案吗?] 返回

- [22] Powles, Julia, and Helen Nissenbaum, 'The Seductive Diversion of "Solving" Bias in Artificial Intelligence', Medium, 2018, <https://medium.com/s/story/the-seductive-diversion-of-solving-bias-in-artificial-intelligence-890df5e5ef53>. [题名中译: “解决”人工智能偏见所造成的诱人转移。] 返回
- [23] Wagner, 'Ethics as an Escape From Regulation'. [题名中译: 伦理作为逃避监管的手段。] 返回
- [24] Whittaker, 'The steep cost of capture'. [题名中译: 俘获的高昂代价。] 返回
- [25] Su et al., 'Critical Affects' and Cath, 'Governing Artificial Intelligence'. [题名中译: 批判性情感; 治理人工智能。] 返回
- [26] Carolyn Ellis, Tony E. Adams, and Arthur P. Bochner, 'Autoethnography: An Overview' *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 12.1 (2011) and Ingo Winkler, 'Doing Autoethnography: Facing Challenges, Taking Choices, Accepting Responsibilities', *Qualitative Inquiry* 24.4 (2018). [题名中译: 自我民族志概述; 做自我民族志: 面对挑战、作出选择、承担责任。] 返回
- [27] Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information*, Cambridge: Harvard University Press, 2016, and O' Neil, Cathy, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, New York: Crown, 2016. [题名中译: 黑箱社会: 控制金钱与信息的秘密算法; 数学毁灭性武器: 大数据如何加剧不平等并威胁民主。] 返回
- [28] Whittaker, 'The Steep Cost of Capture'. [题名中译: 俘获的高昂代价。] 返回
- [29] Daisuke Wakabayashi, 'Big Tech Funds a Think Tank Pushing for Fewer Rules. For Big Tech', *The New York Times*, 24 July 2020. [题名中译: 大型科技企业资助智库, 推动减少对大型科技企业的监管。] 返回
- [30] Cade Metz and Daisuke Wakabayashi, 'Google Researcher Says She Was Fired Over Paper Highlighting Bias in A.I.', *The New York Times*, 3 December 2020. [题名中译: Google研究者称因论文揭示AI偏见而被解雇。] 返回
- [31] Susan Leigh Star, 'This Is Not a Boundary Object: Reflections On the Origin of a Concept', *Science, Technology, & Human Values* 35.5 (2010): 601–17. [题名中译: 这不是边界对象: 对一个概念起源的反思。] 返回
- [32] Phil Agre, 'Your Face is Not a Bar Code: Arguments Against Automatic Face Recognition in Public Places', *Whole Earth* 106.2001 (2001): 74–7. [题名中译: 你的脸不是条形码: 反对在公共场所自动识别人脸的论证。] 返回
- [33] Moten and Harney, 'The University and the Undercommons'. [题名中译: 大学与底层公地。] 返回
- [34] Richard G. Jones, 'Putting Privilege Into Practice Through "Intersectional Reflexivity": Ruminations, Interventions, and Possibilities', in *Reflections: Narratives of Professional Helping* (2010), 其中转引Dwight Conquergood, 'Between Rigor and Relevance: Rethinking Applied Communication,' *Applied Communication In the 21st Century*, London: Routledge, 1995. [题名中译: 通过“交叉性反思”将特权

- 付诸实践：思考、干预与可能性；反思：专业助人工作的叙事；在严谨与相关性之间：重新思考应用传播；21世纪的应用传播。） 返回
- [35] Bellanova, Rocco, Kristina Irion, Katja Lindskov Jacobsen, Francesco Ragazzi, Rune Saugmann, and Lucy Suchman. ‘Toward a Critique of Algorithmic Violence.’ *International Political Sociology* (2021.)15 (1): 121–50. <https://doi.org/10.1093/ips/olab003>. [题名中译：走向对算法暴力的批判。] 返回
- [36] Sotiropoulou and Cranston, ‘Critical Friendship’ . [题名中译：批判性友谊。] 返回
- [37] Alexis Shotwell, *Against Purity: Living Ethically in Compromised Times*, Minneapolis: University of Minnesota Press, 2016. [题名中译：反对纯洁：在妥协时代合乎伦理地生活。] 返回
- [38] Phan et al., ‘Economies of Virtue’ . [题名中译：美德经济。] 返回
- [39] Incite! Women of Color Against Violence (eds), *The revolution will not be funded: Beyond the non-profit industrial complex*, Boston: South End Press, 2007. [题名中译：革命不会得到资助：超越非营利产业复合体。] 返回
- [40] Moten and Harney, ‘The University and the Undercommons’ . [题名中译：大学与底层公地。] 返回
- [41] Honig, ‘A Feminist Theory of Refusal’ . [题名中译：女性主义拒绝理论。] 返回
- [42] Simpson, ‘The Ruse of Consent and the Anatomy of “Refusal” ’ . [题名中译：同意的诡计与“拒绝”的剖析。] 返回
- [43] Abeba Birhane, ‘Algorithmic Injustice: A Relational Ethics Approach’ , *Patterns* 2.2 (2021). [题名中译：算法不公正：一种关系伦理路径。] 返回
- [44] Sotiropoulou and Cranston, ‘Critical Friendship’ . [题名中译：批判性友谊。] 返回
- [45] 参见INCITE!, ‘The Revolution will Not be Funded’ , Whittaker, ‘The Steep Cost of Capture’ , and Ben Tarnoff, *The Making of the Tech Worker Movement*, New York, NY, USA: *Logic* magazine, 2019, <https://logicmag.io/the-making-of-the-tech-worker-movement/>. [题名中译：革命不会得到资助；俘获的高昂代价；科技劳动者运动的形成。] 返回

攫取主义伦理

Sarah Pink

导言

请随我回到开始撰写这篇有关AI、伦理与人的文章的时刻：沿着伦理报告和论文中的链接，我找到了AI4People的伦理框架。这类项目的初衷以社会福祉为本，并积极支持一个重要主张：AI需要受到监管，需要行善。但开始阅读时，我感到不自在；我知道自己即将以人类学家的眼光，批判一个力图使AI合乎伦理的议程的逻辑，而我本该与它站在同一页上（*剧透提醒：请记住“同一页”这个比喻，在我请你点击之前，先别点击链接*）。需要说明的是，我并非专门批评AI4People的伦理框架。这个框架只是一个例子，表明主流话语中通常用来表达人们对于合乎伦理的AI与人之善意的隐喻、叙事和结构，如何暴露出一种与日常现实错位的逻辑。我的议程，以及我的全部工作，都着眼于建立合作伙伴关系，共同走向合乎伦理的未来，而不只是写出一份以指出问题为终点的提纲。但是，要建立合作所需的联系／关系，我们必须让不同方法、学科和逻辑之间的裂缝显现出来。

《AI4People建设美好AI社会的伦理框架：机遇、风险、原则与建议》报告“致力于以确保人们信任、服务公共利益和加强共同社会责任的方式发展AI技术”，并提出了一系列旨在保障这些目标的建议。^[1] 报告作者包括来自哲学、法学和计算机科学等领域、专攻伦理的多位学者及业界人士，却没有人类学、定性社会学等社会科学领域的参与者。从我们目前所处的社会结构来看，报告提出的原则完全合理；但是，它用以表达观念的术语和概念，很难与人类学理解人、伦理和新兴技术的方式相协调。首先，认为AI技术的某些发展可以“确保”人们的信任，就意味着将信任视为一种固定属性，能够从人身上提取并加以捕获。其次，“公共利益”的概念对人作出一维的界定，而不是关注置身于杂乱而充满偶然性的生活之中的真实的人，尽管真实利益与日常伦理恰恰在这种生活中展开。第三，诉诸“社会责任”，使其聚焦于某个社会学的分析单位或层次，而不是生命的经验领域。也就是说，在AI4People的议程中，真实的人的经验与活动显著缺席。

在本章中，我主张采取一种以人为中心的方法，而这种方法必须通过与日常世界中的理论和研究相接触才能显现。我在别处曾把“技术解决主义的伦理方法界定为攫取主义：它们试图识别、捕获人类伦理价值，并将其注入机器，意图让这样的伦理机器产生信任”。^[2] 在此，我进一步展开这一论点，同时提出一个前提：要合乎伦理，AI就不应只是为了人而存在，而应与人共同设计，关注多样性、特殊性和地方性。它不应在设计完成之后才设法确保、获得或赢得人们的信任，或其他任何东西，而应从一开始就在信任关系之中创造出来，关注“日常伦理”。^[3] 我所说的

日常伦理，指的是在日常生活中实际展开、并取决于其具体处境的伦理；伦理并不必然固定到能够在所有情境中一贯适用，而会受到人、物和环境之间关系的细致影响。我的定义建立在现象学人类学对伦理的理解之上，这种理解“揭示了伦理生活是一种以本体论的不确定性与伦理负荷过重为特征的境况”；^[4]在我所实践的人类学中，伦理是不确定的，“是偶然的，涌现于日常世界和生活处境之中”。^[5]这种伦理理解既贯穿于人类学的研究伦理，也贯穿于我们理解他人伦理的方式，因而同样涉及具有反思意识的民族志研究者自身的行动。^[6]

当然，AI4People并不是唯一采取这种路径的议程；它试图回答如何让AI合乎伦理，而其方案与AI当前的发展方式相一致。从某种意义上说，这是朝着伦理方向迈出的一步。但它也体现了科技产业、政府以及工程与计算机科学就如何使AI合乎伦理而提出的主流话语中，那道持续而显眼的缺口。我相信，这些学科和利益相关方都对人与AI抱有共同关切；将人——通过协作式民族志研究实践、接触和干预而相遇的人——带入讨论，可以使我們更好地回应这种关切。这意味着，我们必须确保日常的认识方式与多样性处于伦理思考的最前沿。但只需瞥一眼AI4People报告网页^[7]上的横幅（现在请打开链接），我就确信，这个页面有着明显的性别印记。虽然报告本身由男性和女性共同撰写，但这一事实似乎在横幅的视觉呈现中被抹去了。已经被大量揭露^[8]和批评的“全男小组”（manel，即全由男性组成的讨论小组）^[9]依旧存在，也凸显出我即将讨论的攫取主义伦理问题，同时会揭示技术、数据、伦理与学术界中的性别政治。

攫取主义伦理：开采伦理

我们用以描述数据，以及自动化、互联、智能的新兴技术和系统的隐喻，本身就是争议的场域，因此也构成了可能开展研究和干预的场域。Sally Wyatt指出，作为批判性社会科学研究者，我们需要质疑产业界和政策制定者所采用的攫取性隐喻，这些隐喻将数据界定为一种可以开采的资源。^[10]本章提出一种相关的质疑方式：不替换这些隐喻，而是将它们应用于伦理，以考察在学者已称之为攫取主义的那些数据和AI等新兴技术路径中，伦理处于怎样的位置。^[11]我讨论那些已受到批判性数据研究者批评的类似攫取逻辑，如何被应用于AI伦理，以及这对产业、政策和研究意味着什么。我称之为攫取主义伦理。

那么，攫取主义伦理是如何构成的呢？在一个层面上，如上文的例子所示，伦理成为诱饵，借以从公众或用户那里提取对技术的信任。在另一个层面上，正如我在别处讨论信任与伦理的关系时所提出的，技术解决主义的伦理方法可以被界定为攫取主义，因为“它们试图识别、捕获人类伦理价值，并将其注入机器，意图让这样的伦理机器产生信任”。^[12]这种因果关系是工程科学呈现伦理的典型方式。一个很好的例子，是著名的麻省理工学院“道德机器”实验；^[13]我和其他人曾在别处讨论过它，^[14]但在这里将从一个新的方向展开。

“道德机器”是一款充当调查问卷的游戏，试图从全球数千名参与者那里提取道德判断：参与者要对在线呈现的一系列未来自动驾驶汽车情境作出判断，这些情境以“电车难题”为基础——在这个哲学难题中，游戏参与者必须决定，在事故中，电车应当在若干可能的受害者里撞死谁。游戏被用来从玩家那里提取伦理，或许并不令人意外。Sheila Jasanoff在更广泛地讨论社交媒体时指出，这类技术通过“开采人们的思想、言语、习惯、身体和情感，把它们作为资源，创造新的可销售商品”，从而“从人身上获利”。^[15] Edmond Awad及其同事（全部为男性），与前面讨论的AI4People作者一样，怀有良好意图。他们关注文化差异，并提出：

“我们可以把机器伦理的挑战看作独特的机遇，由共同体一道决定我们认为什么是、什么是错；并确保机器不同于人类，能够毫无差错地遵循这些道德偏好。”

^[16] 再次强调，我并不主要反对这种愿望；我反对的是一种伦理理解，它与全部人类学证据脱节，而这些证据说明，伦理实际上是在日常生活充满偶然性的处境中展开的。正如哲学家Onora O’Neill强调的，用调查或民调量化人类的感受、情感状态和取决于具体情境的决定，存在局限。例如，有关公众信任的民调“无法提供证据，说明人们在针对特定事项决定信任或拒绝信任特定个人或机构时作出了怎样的判断；在这类判断中，他们往往会相当仔细地区分具体情况”。^[17] 同样值得质疑的，是通过人们对不大可能发生的伦理困境的回答而获得的伦理知识的地位：这些回答随后脱离了其来源，而没有被置于伦理实际展开的真实情境之中。

如前所述，把伦理理解为依情境而定，意味着日常伦理是难以捉摸的：它们极难被捕获、被注入组织或机器之中，或被监管。因此，认为可以解决AI嵌入日常生活之后才出现的伦理问题，这种假设的适用范围是有限的。无论是设计伦理机器，还是引入监管与治理，典型的解决方案都在构建风险缓解过程，其逻辑所遵循的因果链，是在日常生活及其伦理之外形成的。相反，当我们把注意力转向日常生活中身为人究竟意味着什么——也就是存在、感受与行动的经验——伦理就无法在日常生活之外被抽象化、固定或预先决定。现象学通过赋予“第一人称和第二人称立场以优先地位”，将我们的注意力引向伦理的主体间性和身体间性。^[18] 正如Mattingly与Throop所言：“伦理远不是一个由文化中表达充分的义务所构成的场域，也不是施加规范性道德秩序以创造顺从主体的场域；学者已经通过经验研究记录了伦理如何提出过度的要求，使生活经验变得陌生而诡异。”^[19] 这意味着，尽管人们不断尝试将伦理抽象为法规，这样的伦理却很可能永远无法契合日常生活的伦理要求。

因此，在本节中，我为法律与科学技术研究学者Sheila Jasanoff所提出的那类问题，给出了两种截然不同的回答。她问道：“在当今复杂的社会中，预见或预先阻止技术的负面影响是谁的责任？我们是否拥有预测和防止伤害所需的工具与手段？”^[20] 麻省理工学院的“道德机器”实验尝试正面回答这个问题：既承担预防负面影响的责任，也创造工具，用以制造合乎伦理的自动驾驶汽车，让人们随后信任并采用它们，由此减少交通死亡和碳排放。但是，Mattingly与Throop提出了另

一个问题，使这种回答及其假定的责任模式都变得复杂：“强调伦理生活的不充分确定性，究竟关涉什么 [……]？将人描述为以潜能或可能性、而不是现实性为特征，意味着什么？声称伦理要求具有一种过量性质、因而无法简化为遵守规定性的规范或规则，又意味着什么？”^[21] 在此情境中，答案是：由于人与自动驾驶汽车之间的关系——进而人与一般AI之间的关系——所呈现的伦理是不确定的，它们既不能外推至机器，也不能被用来预测和防止伤害。（译注：此处原文将 Jasanoff 误拼为 Jasinoff。）

攫取主义伦理与预期性审计：吸引投资

本节考察攫取主义伦理如何参与资本主义的预期愿景。这些愿景是资本主义占据未来、维持其对日常生活的结构性掌控的关键。伦理作为一种可提取的资源，或作为捕获信任的诱饵，可以被注入值得信赖的智能与自动化机器，因而成为人类信任、接受和采用AI这一因果链中的催化剂。在这里，被提取的伦理可以参与创造一种预期性基础设施：在其中，合乎伦理的AI和伦理机器被视为一种技术解决方案，用来应对公众接受自动化技术被认为是难题的情境，^[22] 并由此为那些被设想成社会问题解决方案的技术吸引投资。

为了使AI投资与市场显得可信而现实，攫取主义伦理也与我称为预期性审计的东西相契合。预期性审计属于人类学家早已称之为“审计文化”的一部分。^[23] 审计文化的许多要素本质上具有预期性。以大学伦理委员会为例。^[24] 这个例子我已讨论过多次，但值得在此重提，因为它既与下文所说的学术研究和资助背景相连，也很可能是学术读者亲身经历的一部分。伦理审批通常由制定规则、界定合乎伦理之研究行为的机构级组织（有时是国家级组织）加以管理，其内容包括确保在研究中被认定为不合伦理的事情可能发生的任何风险，都已被预先识别并得到缓解。然而，当伦理成为预期性审计的对象时，唯一能够将伦理纳入核算的方式，就是把它们固定住、捕获下来，再依据伦理法规加以衡量。这样做的结果，是让我们的机构确信，研究实际上会合乎伦理，而且机构已经尽量减少了可能被解释为不合伦理的行为。伦理AI的情形与此类似：AI应当体现的伦理条件通过伦理框架事先规定下来，因此，如同研究者的伦理一样，AI的伦理也可以在它们被放入世界——即日常生活环境——之前接受审计。

举众多例子中的一个：顶尖咨询公司普华永道（PwC）的网站提出了“负责任AI（RAI）”的问题，并声称这是“缓解AI风险的唯一途径”。^[25] 其“负责任AI工具包”包含了对伦理的关注。然而，尽管“把伦理做好就能缓解风险”这种可能性很容易让人信服，它究竟意味着什么，似乎仍悬而未决。2019年一篇国际伦理框架综述发现：“全球正在围绕五项[AI]伦理原则趋于一致（透明、正义与公平、不伤害、责任及隐私），但在如何解释这些原则、为何认为其重要、它们涉及何种问题、领域或行动者，以及应如何落实方面，仍存在实质性分歧。”^[26] 此外，皮尤

研究机构在2021年发布了针对技术专家的调查结果，指出多数专家不认为到2030年，伦理AI设计会被广泛采用。^[27]

不过，皮尤所引述的最重要的评论之一来自danah boyd。她指出，当AI系统与“迷恋效率、规模和自动化”的当代资本主义相一致时，它们就与“增强人的能力、地方情境和包容”这些伦理价值背道而驰。boyd的洞见与上文勾勒的日常伦理相通，后者恰恰强调伦理如何依具体情境而定、具有特殊性。上述观点表明，问题不只是AI的伦理是什么，还包括它从哪里获得伦理，以及这些伦理契合谁的价值观。正如我通过“道德机器”实验所表明的，人们已经尝试通过汇总个体回答，从日常生活中提取伦理；但这类尝试必然无法产生契合日常伦理的伦理，因为它们具有攫取性。我们无法开采伦理。相反，我们必须进入其中，与之共处。这与攫取恰好相反；它要求在具体地方、并伴随着生命持续涌现的过程，进行融合与合作。

实际上，我们作为大学学者所经历的预期性伦理审计，与AI系统受到的审计，其相似之处并不止于共同希望缓解某些行动者“在真实环境中”行动所带来的风险。这两种预期性审计，也都意图核算、规制、控制和缓解如下伦理行为中的风险：一方是AI或研究者这样的主动行动者，另一方则是公众成员、用户、消费者或研究参与者这样的被动行动者。然而，预期性伦理审计与学术研究者或AI系统及技术——那些已经接受审计的人或物——将要被投入其中的日常伦理之间，存在多处错位。在那些预想的单向伦理效应将被激活的日常世界里，实际存在的是非常不同的伦理：Mattingly与Throop所指出的经验性伦理，它们依情境而定、具有脉络性、具身性、主体间性和不确定性。^[28]

资助攫取主义伦理：资助的性别

上文概述了一种观念内在的缺陷：它将人类伦理视为确定的东西，可以从社会中提取，或从专家那里收集，作为社会价值观的代表，再将它捕获并转移到机器之中。^[29]然而，这类伦理方法为一个复杂问题提供了一个（看似）简单的回应：通过一连串具有因果关系的保证，既缓解AI在未来造成伤害的风险，也缓解为创造落实解决方案所需的知识与技术而开展研究时的一组风险。

提出将预先确定的伦理价值嵌入AI的研究，相对而言风险不大，因为它展示了一条清晰的影响路径。当然，它可能还包含其他风险，这些风险与研究希望实现的技术发现所涉及的困难或不确定性有关，那是另一回事。不过，如果你已经相信伦理AI——被注入社会认可的伦理价值的AI——会使人们信任、接受并采用能够造福社会和环境的技术，那么接下来作出这样的假设就不难了：资助旨在制造只会依照可取的人类伦理行动、并由专家批准的伦理框架治理的AI的研究，是一个好主意。因此，攫取主义伦理议程最终会吸引那些怀有善意、参与主流创新议程之叙事和实践的组织与研究。它随后也会支持那些通过此类议程获得资助的学者的职业生涯，并为决定学术成功的研究资助、成果和影响力机制的运转提供润滑。至少从获得资

助的结果来看，性别似乎是支配学术成功的因素之一。2019年，一个由欧盟H2020计划资助、题为“从性别视角看资助分配差异”的项目报告了一系列“无可争议的事实”：在STEM学科和高级学术职位中，女性少于男性；女性获得的科研项目更少、经费更少，评价也更低。^[30]

陷入僵局的攫取主义伦理与日常伦理

本章提出了攫取主义伦理的概念。我认为，它创造了一个范畴，使我们得以揭示并质疑围绕社会中对新兴技术的信任与接受如何形成、以及这些技术的市场如何构成的主流叙事。我探讨了攫取主义伦理与另一个概念——预期性审计——之间的契合。我指出，构造攫取主义伦理与预期性审计的风险缓解范式，也使它们与企业议程和研究议程相契合，因为两者都承诺了产生影响的路径。

那些呼吁伦理框架与监管的伦理AI路径怀有良好意图。Luciano Floridi倡导“伦理优先是为AI制定全球标准的正确方法”，这一点是对的。^[31]然而，要让伦理真正居于首位，还需要更多工作。目前，通过监管自上而下施加伦理的逻辑，与伦理在日常情境中实际发生的方式并不相容，尽管这些情境恰是监管最终试图以合乎伦理的方式影响的对象。日常伦理无法容纳攫取主义伦理在伦理机器问题上所渴求的那种确定性。这在一定程度上呼应了boyd在皮尤调查中的另一项观点：“当我们把伦理想成二元的，认为事物可以是合乎伦理的或不合乎伦理的时，我们就误解了伦理”；这种二元划分确实与“可以把机器制造得合乎伦理”这一想法相一致。同样，把伦理看作可预测因果序列的一部分，也是错误的。伦理无法参与可预测的反应链，仅仅因为它们并非静止不变。这些人类学对伦理的解释，使STEM领域的伦理机器模型及其造福社会的承诺变得复杂。而且，这些伦理理解很难被运用，也很难被教授，因为它们滑动不定、棘手难驯，无法安定下来。除此之外，它们还向那些使STEM在研究中备受重视的社会结构提出另一项挑战，因为正如boyd所说：“在我们能够质疑晚期资本主义的逻辑之前，我们无法有意义地讨论伦理AI。”

日常伦理具有政治性。正如我在本章开头暗示的，这也要求我们关注性别问题。性别具有交叉性，而不是二元的；这意味着，我自己对于全男讨论小组、研究团队以及科研资助结果可能偏向男性的提及，都需要结合其他差异、不平等与不公正形式作更细致的分析。不过，现有证据表明，AI正在一种带有性别印记、往往有利于男性的研发事业中涌现；要向前迈进，我们必须赋予其他声音力量。一个起点，可以是借用《女性主义数据否定宣言》（Feminist Data Manifest-no）^[32]第6点：

“我们拒绝扩张那些将[数据]伦理攫取主义的状态正常化、并主要由货币化和人类经验极端个体化的驱动力所界定的[数据科学]伦理框架形式。”当伦理作为人类经验的一个方面，被从日常生活中提取出来，或被用作诱饵捕获信任等其他日常感受，以构建预期中的市场时，伦理实际上就被商品化了。我们应当像女性主义数

据学者那样，转而“承诺以具有创造性和集体性的生命、生活与造世形式为中心，超越新自由主义逻辑，抵抗由市场驱动、将人类经验商品化的力量”。

致谢

本章最初作为出色的“美德经济”工作坊的一部分构思而成，并从中受益。感谢 Thao Phan、Monique Mann、Jake Goldenfein和Declan Kuch的工作与启发；尤其感谢Ellen Broad、Loren Ruster、Jake Goldenfein和Declan Kuch对本章初稿提出的极富启发性的意见与审阅。

资助披露

本章写作是在我担任澳大利亚研究理事会资助的“自动化决策与社会卓越研究中心”（CE200100005）（2020–2027）首席研究员期间进行的。

参考文献

- Awad, E., Dsouza, S., Kim, R. et al. ‘The Moral Machine Experiment’, *Nature* 563 (2018): 59–64. [题名中译：道德机器实验。]
- Cruz Castro, L. and Sans Menéndez, L. *Literature Review Synthesis Report*. CSIS Institute of Public Goods and Policies, Madrid, 2019. [题名中译：文献综述综合报告。]
- Floridi, L. ‘Establishing the Rules for Building Trustworthy AI’, *Nat Mach Intell* 1 (2019): 261–262. [题名中译：制定构建可信AI的规则。]
- Floridi, L., Cowls, J., Beltrametti, M. et al. AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds & Machines* 28, (2018): 689–707. [题名中译：AI4People：建设美好AI社会的伦理框架——机遇、风险、原则与建议。]
- Jasanoff, S. *The Ethics of Invention*, New York: W. W. Norton & Company, 2016. [题名中译：发明的伦理。]
- Jobin, A., Ienca, M. & Vayena, E. ‘The Global Landscape of AI Ethics Guidelines’, *Nat Mach Intell* 1 (2019): 389–99. [题名中译：AI伦理指南的全球图景。]
- Mattingly, C. and J. Throop. ‘The Anthropology of Ethics and Morality’, *Annual Review of Anthropology* 47:1 (2018): 475–92. [题名中译：伦理与道德的人类学。]
- O’ Nell, O. ‘Accountable Institutions, Trustworthy Cultures’, *Hague J Rule Law* 9 (2017): 401–12. [题名中译：可问责的制度与可信的文化。]
- Pels, P. ‘The Trickster’s Dilemma: Ethics and the Technologies of the Anthropological Self’ in M. Strathern (ed) *Audit Cultures: Anthropological Studies in Accountability*, London: Routledge, 2000, pp. 147–84. [题名中译：诡计者的困境：伦理与人类学自我的技术；审计文化：问责制的人类学研究。]

- Pink, S. 'Ethics in a Changing World: Embracing Uncertainty, Understanding Futures, and Making Responsible Interventions' in xS. Pink, V. Fors, T. O' Dell (eds) *Working in the Between: Theoretical Scholarship and Applied Practice*. Oxford: Berghahn, 2017, pp. 29–51. [题名中译: 变化世界中的伦理: 拥抱不确定性、理解未来并作出负责任的干预; 在其间工作: 理论学术与应用实践。]
- . 'Trust, Ethics and Automation: Anticipatory Imaginaries in Everyday Life' in Pink, S., Lupton, D., Berg, M. and Ruckenstein, M. (eds) *Everyday Automation*, London: Routledge, 2022, 44–58. [题名中译: 信任、伦理与自动化: 日常生活中的预期性想象; 日常自动化。]
- Pink, S., Raats, K., Lindgren T., Osz, K. and Fors, V. 'An Interventional Design Anthropology of Emerging Technologies' in Hojer Bruun, M., Wahlberg, A., Brogaard Kristensen, D., Douglas-Jones, R., Hasse, C., Høyer, K., and Ross Winthereik, B. (eds) *The Handbook for the Anthropology of Technology*, London: Palgrave, 2021, pp. 183–200. [题名中译: 新兴技术的干预性设计人类学; 技术人类学手册。]
- Stilgoe, J., T. Cohen, 'Rejecting Acceptance: Learning from Public Dialogue on Self-Driving Vehicles', *Science and Public Policy*, 48.6 (2021): 849–59. [题名中译: 拒绝接受: 从关于自动驾驶汽车的公众对话中学习。]
- Strathern, M. (ed). *Audit Cultures: Anthropological Studies in Accountability*, London: Routledge, 2000. [题名中译: 审计文化: 问责制的人类学研究。]
- Wyatt, S. 'Metaphors in Critical Internet and Digital Media Studies', *New Media & Society*, 23.2 (2021): 406–16. [题名中译: 批判性互联网与数字媒体研究中的隐喻。]

原注

- [1] L. Floridi, J. Cows, M. Beltrametti, et al. 'AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations', *Minds & Machines* 28, (2018): 22, https://www.eismd.eu/wp-content/uploads/2019/11/AI4People%E2%80%99s-Ethical-Framework-for-a-Good-AI-Society_compressed.pdf. [题名中译: AI4People: 建设美好AI社会的伦理框架——机遇、风险、原则与建议。] 返回
- [2] S. Pink, 'Trust, Ethics and Automation: Anticipatory Imaginaries in Everyday Life' in S. Pink, D. Lupton, M. Berg & M. Ruckenstein (eds) *Everyday Automation*, London: Routledge, 2022. [题名中译: 信任、伦理与自动化: 日常生活中的预期性想象; 日常自动化。] 返回
- [3] Pink, 'Trust, Ethics and Automation'. [题名中译: 《信任、伦理与自动化》。] 返回
- [4] C. Mattingly and J. Throop, 'The Anthropology of Ethics and Morality', *Annual Review of Anthropology* 47.1 (2018): 483. [题名中译: 伦理与道德的人类学。] 返回
- [5] Pink, 'Trust, Ethics and Automation'. [题名中译: 《信任、伦理与自动化》。] 返回
- [6] M. Strathern (ed) *Audit Cultures: Anthropological Studies in Accountability*, London: Routledge, 2000; P. Pels, 'The Trickster' s Dilemma: Ethics and the Technologies of

- the Anthropological Self' in M. Strathern (ed) *Audit Cultures*. [题名中译：审计文化：问责制的人类学研究；诡计者的困境：伦理与人类学自我的技术；审计文化。] 返回
- [7] <https://www.eismd.eu/featured/ai4peoples-ethical-framework-for-a-good-ai-society/>. 返回
- [8] @allmalepanels, <https://twitter.com/allmalepanels?lang=en>. 返回
- [9] J.K. Rodriguez and E.A. Guenther, 'What' s Wrong With "Manels" and what Can We Do About Them?' The Conversation, 15 October 2020, <https://theconversation.com/whats-wrong-with-manels-and-what-can-we-do-about-them-148068>. [题名中译：“全男小组”有什么问题，我们能做些什么？] 返回
- [10] S. Wyatt, 'Metaphors in Critical Internet and Digital Media Studies' , *New Media & Society*, 23.2 (2021): 406–16. [题名中译：批判性互联网与数字媒体研究中的隐喻。] 返回
- [11] S. Jasanoff, *The Ethics of Invention*. New York: W. W. Norton & Company, 2016; Wyatt, 'Metaphors' . [题名中译：发明的伦理；隐喻。] 返回
- [12] Pink, 'Trust, Ethics and Automation' . [题名中译：《信任、伦理与自动化》。] 返回
- [13] E. Awad, S. Dsouza, R. Kim et al., 'The Moral Machine Experiment' , *Nature*, 563 (2018): 59–64. 另见： <https://www.moralmachine.net/>. [题名中译：道德机器实验。] 返回
- [14] S. Pink, K. Raats, T. Lindgren, K. Osz, and V. Fors, 'An Interventional Design Anthropology of Emerging Technologies' in Maja Hojer Bruun, Ayo Wahlberg, Dorthe Brogaard Kristensen, Rachel Douglas-Jones, Cathrine Hasse, Klaus Høyer, and Brit Ross Winthereik (eds) *The Handbook for the Anthropology of Technology*, London: Palgrave, 2021; Pink, 'Trust, Ethics and Automation' . [题名中译：新兴技术的干预性设计人类学；技术人类学手册；信任、伦理与自动化。] 返回
- [15] Jasanoff, 259. 返回
- [16] Awad et al., 63. 返回
- [17] O. O' Neill, 'Accountable Institutions, Trustworthy Cultures' , *Hague J Rule Law* 9 (2017): 405. [题名中译：可问责的制度与可信的文化。] 返回
- [18] Mattingly and Throop, 483. 返回
- [19] Mattingly and Throop, 485-6. 返回
- [20] Jasanoff, 7. 返回
- [21] Mattingly and Throop, 'The Anthropology of Ethics and Morality' , 486. [题名中译：伦理与道德的人类学。] 返回
- [22] J. Stilgoe, T. Cohen, 'Rejecting Acceptance: Learning from Public Dialogue on Self-Driving Vehicles' , *Science and Public Policy*, 48.6 (2021): 849–59. [题名中译：拒绝接受：从关于自动驾驶汽车的公众对话中学习。] 返回
- [23] 例如, Strathern, *Audit Cultures* (《审计文化》) 。 返回

- [24] S. Pink, ‘Ethics in a Changing World: Embracing Uncertainty, Understanding Futures, and Making Responsible Interventions’ (《变化世界中的伦理：拥抱不确定性、理解未来并作出负责任的干预》), 载S. Pink、V. Fors、T. O’ Dell编, *Working in the Between: Theoretical Scholarship and Applied Practice* (《在其间工作：理论学术与应用实践》), Oxford: Berghahn, 2017。需要说明的是, 我支持伦理审查程序, 因为当它们得到良好实施时, 既会促使人反思, 也会促进合乎伦理的行为。返回
- [25] PwC, <https://www.pwc.com/gx/en/issues/data-and-analytics/artificial-intelligence/what-is-responsible-ai.html>. 返回
- [26] A. Jobin, M. Ienca, and E. Vayena, ‘The Global Landscape of AI Ethics Guidelines’, *Nat Mach Intell* 1 (2019): 389. [题名中译: AI伦理指南的全球图景。] 返回
- [27] <https://www.pewresearch.org/internet/2021/06/16/experts-doubt-ethical-ai-design-will-be-broadly-adopted-as-the-norm-within-the-next-decade/>. 返回
- [28] Mattingly and Throop, ‘The Anthropology of Ethics and Morality’. [题名中译: 伦理与道德的人类学。] 返回
- [29] 本章专门针对制造伦理机器的挑战提出这一问题。它还引出了更广泛的问题, 即这种立场如何与伦理、监管和治理在社会中的地位相协调; 这些问题不在本章的讨论范围内。返回
- [30] L. Cruz Castro and L. Sans Menéndez, *Literature Review Synthesis Report*. CSIS Institute of Public Goods and Policies, Madrid, 2019. [题名中译: 文献综述综合报告。] 返回
- [31] L. Floridi, ‘Establishing the Rules for Building Trustworthy AI’, *Nat Mach Intell* 1 (2019): 262. [题名中译: 制定构建可信AI的规则。] 返回
- [32] Feminist Data Manifest-No (《女性主义数据否定宣言》), <https://www.manifestno.com/home>; 《女性主义数据宣言》第6点写道: “我们拒绝扩张那些将数据攫取主义状态正常化、并主要由货币化和人类经验极端个体化的驱动力所界定的数据科学形式。我们承诺以具有创造性和集体性的生命、生活与造世形式为中心, 超越新自由主义逻辑, 抵抗由市场驱动、将人类经验商品化的力量。” 返回

“伦理AI”的发明：大型科技企业如何操纵学术界以规避监管

Rodrigo Ochigame

本文最初于2019年12月20日发表于 *The Intercept*，获作者许可在此转载。

笼罩着麻省理工学院媒体实验室前主任伊藤穰一（Joichi Ito）的伦理丑闻，其讽刺之处在于，他过去曾领导有关伦理的学术项目。在他与金融家 Jeffrey Epstein 的资金关系曝光后，伊藤辞去了在麻省理工学院的多项职务、哈佛法学院的客座教授职位，以及 John D. and Catherine T. MacArthur Foundation、John S. and James L. Knight Foundation 和纽约时报公司的董事职务。Epstein 被控贩运未成年少女进行性交易，其中最小的只有14岁。

许多旁观者不解，伊藤为何能作为人工智能伦理学者发挥如此大的影响。实际上，他推动的项目对于确立“伦理AI”话语至关重要，而如今这种话语在学术界与主流媒体中无处不在。2016年，时任总统 Barack Obama 称他为AI与伦理方面的“专家”。自2017年起，伊藤通过规模为2700万美元的人工智能伦理与治理基金资助了许多项目；该基金以麻省理工学院媒体实验室和哈佛大学伯克曼·克莱因互联网与社会研究中心为依托。这一切关于“伦理”的讨论，究竟是怎么回事？

我曾在媒体实验室伊藤的AI伦理小组担任研究生研究员，为期14个月。我于2019年8月15日停止了这项工作，就在伊藤首次就自己与 Epstein 的关系发表“道歉”之后。在道歉中，他承认既为媒体实验室，也为自己在外部的风险投资基金接受了这位金融家的钱。他没有披露，在这些资金往来发生时，Epstein 已在佛罗里达州对一项涉及儿童卖淫的指控认罪；他也没有披露自己采取了许多措施，让 Epstein 的名字不出现在正式记录中——这些后来才由《纽约客》揭露出来。

受到举报人 Signe Swenson 及其他发声者的鼓舞，我决定报道自己所了解到的、伊藤在塑造AI伦理领域中所扮演的角色，因为这事关公共利益。这一领域的出现是近来的现象：过去，AI研究者大多对伦理研究不感兴趣。一位曾在媒体实验室工作的同事回忆说，麻省理工学院已故AI先驱 Marvin Minsky 常说：“伦理学者，就是不管你心里想什么都要挑出问题的人。”^[1] 那么，AI研究者为什么突然开始谈论伦理了？

在媒体实验室，我了解到，主要由伊藤倡导的“伦理AI”话语，在策略上配合着硅谷的一项行动，其目的在于避免有争议的技术受到具有法律强制力的限制。参与这一行动的关键团体之一，媒体实验室也是其成员，在加利福尼亚州提出了政策建

议；这些建议与我同几位实验室同事开展的研究所得结论相矛盾。我们的研究使我们反对在决定是否于审判前羁押某人时使用计算机算法。伊藤本人后来在与金融和科技界高管的私下会议中也抱怨，该团体的建议无异于对棘手伦理问题的“粉饰”。他向一位亿万富翁吐露，在羁押决定中，“我们试图阻止使用那些看起来不太管用的算法，而他们却把我们想说的话冲淡了”。

我还亲眼看到麻省理工学院帮助美国军方把无人机战争的道德复杂性搁置一旁：学校举办了一场由前国务卿、臭名昭著的战争罪犯 Henry Kissinger 主讲的、关于AI与伦理的肤浅演讲，并为美国国防部用于战争的“AI伦理原则”提供意见。这些原则接受“可被允许的有偏见”算法，并避而不用“公平”一词，因为五角大楼认为“战斗不应该公平”。

伊藤没有回应置评请求。

在科技行业面对越来越多批评及要求法律监管的呼声之际，麻省理工学院为“大型科技企业能够自行约束其人工智能使用”这一观念增添了可信度。仅在2018年，就发生了多起争议事件：3月曝光，Facebook 将超过5000万用户的私人数据泄露给一家受 Donald Trump 总统竞选团队聘用的政治营销公司；同月曝光，Google 与五角大楼签约，提供在战区使用的计算机视觉软件；5月曝光，Amazon 向警察部门出售人脸识别技术；6月曝光，Microsoft 与美国移民及海关执法局签订合同；9月曝光，IBM 与纽约市警察局秘密合作，对监控视频中的人进行人脸识别和种族分类。在 #TechWontBuildIt（科技工作者拒绝制造这些技术）的口号下，这些公司数千名员工组织抗议、传递请愿书，反对上述合同。从 #NoTechForICE 到 #Data4BlackLives，多场草根运动要求以法律限制计算技术的某些用途，例如禁止警方使用人脸识别。

与此同时，企业试图转移讨论焦点，转而强调自愿的“伦理原则”、“负责任的实践”，以及用“偏见”和“公平”来表述的技术调整或“保障措施”，例如要求或鼓励警方采用“无偏见”或“公平”的人脸识别。2018年1月，Microsoft 发布了AI“伦理原则”，第一项就是“公平”。5月，Facebook 宣布“致力于合乎伦理地开发和部署AI”，并推出名为“Fairness Flow”的“寻找偏见”工具。6月，Google 公布了AI研发的“负责任实践”。9月，IBM 宣布推出名为“AI Fairness 360”的工具，旨在“检查数据集和机器学习模型中不受欢迎的偏见”。2019年1月，Facebook 拨款750万美元，在德国慕尼黑建立AI伦理中心。3月，Amazon 与美国国家科学基金会共同资助了一项2000万美元的“AI公平性”计划。4月，Google 取消了其AI伦理委员会，原因是委员会选入了右翼传统基金会主席 Kay Coles James；她公开反对跨性别者，这一任命引起强烈反弹。这些企业项目经常援引伊藤曾通过麻省理工学院—哈佛基金至少部分资助过的学术研究。

要说明企业的议程，有必要区分针对某项技术的三种监管可能性：（1）完全没有法律监管，“伦理原则”和“负责任实践”只靠自愿；（2）适度的法律监管，鼓励或要求进行不会与利润发生重大冲突的技术调整；或（3）限制性的法律监管，抑制或禁止部署该项技术。不出所料，科技行业倾向于支持前两种、反对最后一种。企业资助的“伦理AI”话语恰恰让这种立场得以成立。以人脸识别为例。今年，加利福尼亚州的旧金山、奥克兰和伯克利，以及马萨诸塞州萨默维尔的市级立法机关，都通过了对人脸识别技术的严格禁令。与此同时，Microsoft 却游说支持限制较少的立法，要求进行“偏见”测试之类的技术调整，其中最显著的案例发生在华盛顿州。一些大公司甚至可能比起完全没有法律监管，更喜欢这种温和监管，因为大公司更容易投入专门团队来开发符合法规要求的系统。

因此，硅谷大力推销“伦理AI”，构成了一场策略性的游说行动，并将学术界纳入其中，为这场行动赋予正当性。伊藤在这种企业与学术界的亲密往来中发挥了关键作用，经常与科技公司高管会面。麻省理工学院—哈佛基金的首任主任，曾是 Google AI 领域的“全球公共政策负责人”。伊藤及其同事通过该基金资助了许多项目，包括创办计算机科学领域一个著名的“公平、问责与透明”会议；该会议的其他赞助者包括 Google、Facebook 和 Microsoft。

尽管硅谷的游说行动自2016年起凝聚了学术界对“伦理AI”和“公平算法”的兴趣，但更早以前已经有少量论文讨论这些议题，只是采用的框架有所不同。例如，Microsoft 的计算机科学家于2012年发表的论文，可以说开创了“算法公平性”这一领域。2016年，该论文第一作者 Cynthia Dwork 成为哈佛大学计算机科学教授，同时在哈佛法学院与 Microsoft 任职。2017年，我修读她在哈佛讲授的密码学与统计学数学基础课程时，访谈了她，并问她为何对研究公平的算法定义产生兴趣。据她所述，她个人长期关注歧视性广告问题，但 Microsoft 管理者鼓励她沿着这条路线开展工作，是因为公司当时正在开发一个新的在线广告系统，而提供一项“不存在监管问题”的服务会带来经济优势。（公平地说，我相信，尽管她的理念遭到了企业俘获，Dwork 的个人意图仍是真诚的。Microsoft 拒绝就本文发表评论。）

在麻省理工学院和哈佛迈出最初几步后，许多其他大学及新成立的研究机构也收到科技行业的资金，用于AI伦理研究。这类组织大多也由科技公司的现任或前任高管领导。例如，Data & Society Research Institute 由一名 Microsoft 研究员领导，最初获得 Microsoft 的拨款；纽约大学 AI Now Institute 由另一名 Microsoft 研究员共同创立，部分经费来自 Microsoft、Google 和 DeepMind；斯坦福以人为本人工智能研究院由一名前 Google 副总裁担任共同主任；加州大学伯克利分校数据科学部由一位 Microsoft 资深人士领导；麻省理工学院施瓦茨曼计算学院则由 Amazon 董事会成员领导。我在媒体实验室期间，伊藤与所有这些组织的高管和筹划者都保持着频繁联系。

大型科技企业的资金与主导作用，事实证明与对伦理的诚实探究并不相容——至少从我与“造福人类与社会的人工智能伙伴关系”（Partnership on AI to Benefit People and Society, PAI）打交道的经历来看是这样。该团体于2016年由Microsoft、Google/DeepMind、Facebook、IBM和Amazon创立。媒体实验室也是PAI的成员；PAI自称是“多利益相关方机构”，并宣称自己“不是游说组织”。在2018年4月美国众议院监督与政府改革委员会的一场听证会上，其执行主任声称，该组织仅仅是“政策制定者可以利用的资源——例如，开展研究，为AI最佳实践提供信息，探索某些AI系统的社会后果，以及围绕AI系统开发和使用的政策”。

不过，即使该伙伴关系的活动未必达到必须登记为游说者的法定门槛——例如，试图直接影响个别民选官员的投票——它无疑一直在寻求影响立法。例如，2018年11月，伙伴关系的工作人员邀请学术成员参与一份集体声明，针对一项刑罚改革的参议院法案（S.B. 10），向加利福尼亚州司法委员会提出意见。该法案在取消现金保释的同时，扩大了算法风险评估在审前决策中的使用，并要求司法委员会“处理评估工具中任何隐含偏见的识别与缓解问题”。伙伴关系的工作人员写道：“我们相信，有空间影响这项立法，以及更广泛的CJS [刑事司法系统] 应用。”

2018年12月，我与媒体实验室的三位同事，对伙伴关系试图影响立法的举措提出了严正异议。我们注意到，该组织的政策建议始终与其企业议程一致。在刑罚问题上，我们的研究使我们坚决反对采用风险评估工具，也拒绝那些据称能令工具变得“无偏见”或“公平”的技术调整建议。但伙伴关系的声明草案似乎正如一位同事在发给伊藤等人的内部邮件中所说：“通过把这个问题强调为技术问题，因而可以借助更好的数据集等方法解决，来认可风险评估（RA）的使用。”第二位同事也认为：“PAI的声明很软弱，恰恰可能做出我们一直警告要警惕的事情，即通过这些由行业主导的监管举措获得正当性。”第三位同事写道：“就刑事司法方面的工作而言，PAI在这一领域的作为相当令人警惕，而且在我看来存在严重误导。我同意Rodrigo的看法：PAI与ACLU [美国公民自由联盟]、MIT和其他学术/非营利机构的关联，实际上最终发挥了赋予正当性的作用。无论ACLU、MIT，还是任何非营利组织，在PAI内部都没有任何权力。”

更糟的是，伙伴关系的建议似乎与一个反对扩大监狱规模的草根组织联盟所开展的工作不相吻合。该联盟包括“黑人的命也是命”运动、监狱废除主义团体Critical Resistance（我曾在那里做志愿者），以及由无证移民和酷儿/跨性别青年领导的Immigrant Youth Coalition。草根联盟指出：“认为任何风险评估工具都能够处理偏见的观点，忽视了当前与过去警务实践中的种族差异。”支持这一主张的理论与实证理由非常丰富，因为风险评估通常建立在逮捕、定罪或监禁数据之上，而所有这些数据都无法良好地代表个人行为或倾向。联盟继续说道：“归根结底，风险评估工具造成了种族定性、审前羁押和定罪的反馈循环。一个人的自由不应被化约

为某种算法。”相比之下，伙伴关系的声明聚焦于“负责任部署的最低要求”，涵盖的议题包括“有效性WithData抽样偏差、统计预测中的偏差；适当预测目标的选择；人机交互问题；用户培训；政策与治理；透明度与审查；可复现性、流程和记录保存；以及部署后的评估”。

诚然，伙伴关系的工作人员回应了对草案的批评，在声明最终版本中指出：“在PAI成员及更广泛的AI社群中，许多专家进一步提出，尚若不举行个体化听证，就仅凭风险评估分数拘留个人，永远都不可能是公正的。”这一温顺的让步，承认了现在或许还不是单靠软件、不经法官介入或任何其他“个体化”司法程序就开始把人送进监狱的时候。这种让步之所以容易作出，是因为伙伴关系中的主要公司都不出售用于审前决策的风险评估工具：这种技术不仅争议太大，市场也太小。^[2]

2018年12月，我和同事们敦促伊藤退出伙伴关系。我主张：“如果学术和非营利组织想要有所作为，唯一可行的策略就是退出PAI，发表公开声明，并组成一个与之抗衡的联盟。”随后一位同事提议：“在刑事司法预测分析这一领域，许多其他组织正在从事更实质、更具有变革性的工作——把目前用于支持PAI的资金拿来支持他们的工作，会怎么样？”我们相信伊藤有足够的自主权这样做，因为麻省理工学院—哈佛基金主要由奈特基金会支持，尽管大部分资金来自科技投资者：eBay创始人 Pierre Omidyar 通过 Omidyar Network 提供的资金，以及 LinkedIn 联合创始人、Microsoft 董事 Reid Hoffman 的资金。我写道：“如果来自非营利基金会和个人捐赠者的数千万美元，都不足以让我们采取大胆立场、站到正确的一边，那么我不知道还有什么能做到。”^[3]

伊藤确实承认了这个问题。他刚刚收到对冲基金 Two Sigma 联合主席、麻省理工学院校董会成员 David M. Siegel 的消息。Siegel 提议，为硅谷“搜索和社交媒体”公司建立一个自律结构，以金融业监管局（Financial Industry Regulatory Authority, FINRA）为模型；FINRA是一家私人法人机构，充当华尔街证券公司的自律组织。伊藤回应 Siegel 的提议说：“我觉得公民社会在这些行业团体中没有得到很好的代表。我们一直在参与人工智能伙伴关系，但他们会冲淡我们想表达的内容；我们试图阻止使用那些看起来不太管用的算法，例如用于审前保释的风险评分。我想，在个人数据和社交媒体方面，我对自律存在担忧。例如，一场彻底的种族灭绝[针对罗兴亚人，一个缅甸以穆斯林为主的少数民族群]就是利用 WhatsApp 发生的，而 Facebook 知道这件事正在发生。”^[4]

然而，企业与学术界的联盟太过牢固，也太过便利。媒体实验室留在了伙伴关系中，伊藤继续与硅谷和华尔街的高管及投资者亲密交往。他把亿万富翁 Siegel 称作“潜在资助者”。我看到，在与这些人交往时，伊藤经常表达对其业务的道德忧虑，但方式很友好，因为与此同时，他正在为MIT或者自己的风险投资基金向他们要钱。对企业—学术界的“伦理学者”而言，友善的批评可以成为建立商业关系的

筹码。Siegel 回复伊藤：“我很乐意与你更多地讨论这个话题。FINRA不是行业团体，它只是由行业出资。见面时我会进一步解释。我赞同你的担忧。”

在私下会议中，伊藤与科技公司高管相当坦率地讨论企业游说。一月，我和同事们参加了一场与 Mustafa Suleyman 的会议；他是伙伴关系的创始联合主席，也是AI初创公司 DeepMind 的联合创始人，Google 于2014年以约5亿美元收购了该公司。会上，伊藤和 Suleyman 谈到，推广“AI伦理”已变成一项“粉饰”工程，尽管他们声称自己最初的意图更为高尚。在筹备会议的消息中，伊藤对我和同事们写道：“不过，我确实知道，因为在 Mustafa 成立PAI时曾与他交谈，他本来是希望这个团体能更有实质作用，而不仅仅是‘粉饰’。我想它只是沿着这种事情通常会有轨迹发展了。”Suleyman 没有回应置评请求。

不论个别行动者的意图如何，企业游说集团塑造学术研究的努力极为成功。如今以“AI伦理”为名的研究数量庞大。公平地说，其中有些研究有用而且细致，尤其是人文与社会科学领域的研究。然而，经费充足的“伦理AI”研究，大多与科技游说集团的议程一致：自愿地或适度地调整有争议技术的部署，而非以法律加以限制。五家公司，仅仅动用其预算的一小部分，怎么就能如此迅速地影响并框定这么多学科中的这么多学术活动？伊藤没有受过正式训练，却被定位为AI伦理“专家”，而这个领域在2017年以前几乎并不存在，这很奇怪。但更奇怪的是，两年之后，成熟学科中备受尊敬的学者，竟然需要证明自己与一个由企业游说集团凭空制造的领域有何关联。

这个领域也与美国军方发生了关系：不仅体现在官方如何回应围绕定点杀戮技术的道德担忧，也体现在硅谷公司之间对利润丰厚的军事合同的争夺中。2019年11月1日，国防部创新委员会发布了有关“AI伦理原则”的建议。该委员会主席 Eric Schmidt 在2016年时任 Alphabet (Google 的母公司) 执行董事长；正是在那一年，奥巴马政府的国防部长 Ashton B. Carter 成立了该委员会并任命了他。据 ProPublica 报道：“Schmidt 在 Carter 任内的影响力已经很大，而当 [James] Mattis 成为 [Trump 政府的] 国防部长后，他的影响力更大了。”委员会中有多位 Google、Microsoft 和 Facebook 的高管，引发了有关利益冲突的争议。一名负责监督利益冲突的五角大楼雇员，在质疑“五角大楼不仅与 [Amazon 首席执行官 Jeff] Bezos，而且与 Google 的 Eric Schmidt 关系过于亲密”后，被调离创新委员会。这种关系可能为大型科技企业带来丰厚收益：AI伦理建议发布前不到一周，五角大楼刚把一份价值100亿美元的云计算合同授予 Microsoft，而 Amazon 正就此提出法律挑战。

这些建议试图推动五角大楼增加对AI的军事投资，并采用硅谷公司开发、出售的那类“伦理AI”系统。创新委员会称五角大楼是一个“高度合乎伦理的组织”，并提出将其“既有伦理框架”扩展到AI。为此，委员会援引 Google、Microsoft 和 IBM 的AI伦理研究小组，以及由麻省理工学院—哈佛基金资助的学者。不过，其中存在

保留。例如，委员会指出，虽然“‘公平’一词经常被AI社群提及”，这些建议却避开这个词，因为“国防部的口号是战斗不应公平，国防部的目标是创造条件，对任何潜在对手维持不公平优势”。因此，“某些应用将具有可被允许、可获正当辩护的偏向”，具体而言，就是“更成功地瞄准某些敌对战斗人员”。五角大楼的AI伦理观，将许多重要的道德审议可能性排除在外，例如禁止利用无人机实施定点杀戮。

来自企业、学术界和军方的“伦理AI”倡导者，为了共同利益展开了密切合作。例如，伊藤告诉我，他曾非正式地建议 Schmidt，其私人基金会应当资助哪些学术界AI伦理学者。有一次，伊藤甚至来征询我的间接意见，问 Schmidt 是否应该资助某位教授；后来，该教授与伊藤一样，担任了五角大楼创新委员会的“专家顾问”。2019年2月，伊藤与 Carter 一同参加了题为“为人民而计算：伦理与AI”的专题讨论，参加者还包括 Microsoft 和 Google 的现任及前任高管。这场讨论属于麻省理工学院价值10亿美元、专门研究AI的新学院成立庆典的一部分。庆典的其他发言者包括：Schmidt，谈“为市场而计算”；Siegel，谈“我如何学会停止忧虑并爱上算法”；以及 Henry Kissinger，谈“启蒙如何终结”。当 Kissinger 宣称可能出现“一个依靠由数据和算法驱动的机器、却不受伦理或哲学规范约束的世界”时，MIT礼堂外的抗议活动则呼吁人们关注他在越南、柬埔寨和老挝的战争罪行，以及他对其他地区战争罪行的支持。在自动化瞄准的时代，美国军方又会把哪些暴行辩称为受到“伦理”规范约束，或者辩称它们由超出人类能动性、与罪责范围的机器执行？

任何站得住脚的“伦理”主张，都不能回避以具有法律强制力的限制来约束大规模监控和系统性暴力技术部署的迫切性。在这样的限制建立之前，关于计算的道德与政治审议，将始终从属于媒体实验室口号“部署，否则就死”（Deploy or Die）所表达的逐利命令。当一些人部署技术时，即使表面上“合乎伦理”，另一些人也在死去。

资助披露

作者于2018年7月至2019年8月在麻省理工学院媒体实验室担任研究生研究员。撰写本报告时，作者获得麻省理工学院历史、人类学与科学、技术及社会项目的博士奖学金，以及 Jain Family Institute 的研究奖学金（资助的是另一项社会科学工作）。作者谢绝了 *The Intercept* 的稿酬。

原注

[1] 在最近解封的法庭文件中，受害者 Virginia Roberts Giuffre 作证称，Epstein 指使她与 Minsky 发生性关系。返回

[2] 另一方面，人脸识别技术的市场大得多，Microsoft、Google、Facebook、IBM 和 Amazon 都在其中经营。 [返回](#)

[3] Omidyar 为 *The Intercept* 提供资金。 [返回](#)

[4] Facebook 已承认，其平台被用于煽动缅甸的暴力；新闻报道记录了 Facebook 平台上的内容如何助长该国的种族灭绝，尽管人权活动人士与研究者曾反复向 Facebook 高管发出警告。在有报道指出 WhatsApp 被用于在印度选举期间传播虚假信息后，Facebook 旗下的短信服务 WhatsApp 增加了用户转发消息的难度。 [返回](#)

第二部分：场域

生态灭绝并不合乎伦理：政治生态学与资本主义AI伦理

Sy Taffel、Laura Bedford 与 Monique Mann

引言

围绕与数据和数字技术相关的新型攫取主义而出现的话语，将数据描述为“新石油”、^[1]驱动信息经济的关键原材料、^[2]一种新的流通资本，^[3]或一种新的殖民主义。^[4]虽然表述框架各异，但这些关于数字技术文化的概念化，都揭示了诸多伤害：它们加剧既有社会不平等、^[5]歧视某些群体、^[6]并实施预测性的控制逻辑。^[7]然而，这些由数据驱动的“新”攫取形式，与“旧”攫取形式并非彼此分离；后者涉及获取数字技术及其基础设施所需的材料和能源，也涉及21世纪数据处理、存储和传输所导致的生产、消费与废弃物大幅增长带来的社会和环境的影响。

对于计算系统相关伤害的回应，往往被置于“AI伦理”领域之内。^[8]在这一领域，增强隐私保护，或让AI更加“公平”、“可问责”和“透明”的关注重点，遮蔽了更广泛的伦理问题：AI的生产、维护、使用和废弃处理所需的能源、劳动与材料，对当代及未来生态系统都造成了重大伤害。构成当代AI的代码、碳、钴、铜及众多其他材料所组成的行星尺度组装体，在气候变化以及“资本世”

(Capitalocene) 的相关生态危机中扮演着重要角色。^[9]鉴于《巴黎协定》的目标是将全球升温控制在比工业化前水平高不到2摄氏度，最好不超过1.5度，^[10]任何有意义的AI伦理讨论，都必须处理比目前广泛得多的问题。这包括承认AI对气候变化影响的助长作用：干旱、生物多样性丧失、洪水、物种灭绝、沿海地区被淹，以及粮食和水安全问题，此外还有许多其他物质后果。我们认为，在考虑AI时不处理这些伤害，本身就是不合乎伦理的。本章借助政治生态学领域，探讨要让AI能够被视为“合乎伦理”，究竟需要怎样的系统性与结构性变革。

我们主张，当前全球经济安排具有生态灭绝倾向，而数字技术和AI在其中被定位为经济增长的关键驱动力，同时也被定位为气候危机本身的解决方案，^[11]这就要求重新界定AI伦理。在把我们带到全球生态危机这一境地的过程中，技术不能被理解为中立的；更进一步，在我们应对危机时，它的角色也不会是中立的。因此，在预先构想未来全球安排时，必须把技术的角色置于中心。政治生态学致力于揭示环境危机的政治、社会与经济维度，因为这些危机往往被错误地视为“自然的”、“非政治的”。我们借鉴的是一批日益增长的文献，它们把这一概念视角集中运用于技术、^[12]数据、^[13]数字经济^[14]和电子废弃物。^[15]

这些近期研究运用政治生态学框架，展示数字技术的生态与政治经济维度；数字技术通常被错误地认定为“去物质化的”或“人工的”。这两个词都有效地隐藏了这些人工物的物质维度，使它们经常被描绘成非政治的、中立的数学对象。尽管在想象中的自然／文化二元论中，生态与技术的位置截然不同，它们却都有一种共同倾向：被定位为政治中立的问题，应该交给不同科学分支处理，而不是被理解为与权力系统纠缠在一起。对权力、不平等、政治和集体行动的关注，以政治生态学的伦理观为基础：在规范意义上，伦理关乎通过社会行动，把权力重新分配给从属群体。^[16]这使政治生态学的方法与正在出现的主张相呼应：AI话语应当“超越‘伦理’的语言，直面权力与政治经济”。^[17]这里的关键在于，“AI伦理”常常成为强大企业规避监管的方法：企业采取自愿实践，聚焦技术修补，却未能有意义地处理AI相关伤害。

虽然已有一些学术研究拓宽AI伦理，将AI的生态影响纳入其中，^[18]但这些贡献往往局限于狭窄的技术考量。因此，我们首先对比两种进路：一种是以生态伤害和不平等权力关系为中心的政治生态学进路；另一种是把关注焦点放在训练机器学习（ML）模型的碳成本上的“绿色AI”话语。^[19]AI的生态固然包括进行机器学习训练与推理的数据中心，但我们认为，其影响远远超出数据中心；应把这些设施置于AI系统运作所需的能源、劳动、知识及无数材料的流动之中。不过，我们对AI生态政治影响的考察仍从数据中心开始，因为它既是可持续AI文献的核心焦点，也是数字基础设施批判研究的关键场所。^[20]

随后，我们批评两种突出而相互关联的话语。它们都暗示，面对生态危机，数字技术能够维持现有社会经济秩序。第一种话语认为，AI促进一种超高效率的生产模式，使财富生产与物质约束脱钩，从而让经济照常运转。我们展示，透过政治生态学视角，这套推理经不起推敲。我们批评的第二种话语，则建立在一个错误主张之上：技术创新会直接解决当代生态危机。我们质疑那些把资本主义危机重塑为创新科技企业商业机会的技术解决主义回应。相反，我们认为，围绕AI伦理的霸权话语为生态灭绝提供正当性，同时积极阻碍必要的、系统性和结构性的社会、政治与文化转型。例如，这种转型包括在国内及国际刑法中承认生态灭绝。^[21]

最后，我们简要勾勒超高效率绿色增长和技术解决主义之外的替代道路。我们承认，在政治生态学研究内部，对于应当追求怎样的可能世界、以及如何抵达那里，存在大量争论。这些争论也延伸到技术应在转型过程及预先构想的社会中发挥何种作用。这些立场包括去增长、生态社会主义、共生性、共用实践，以及数字基础设施的公共服务模式。本章篇幅有限，无法深入讨论所有这些观点。不过，指出这些替代方案，可以勾勒通往更合乎伦理的AI形式的后资本主义道路。强调既有经济实践所造成生态伤害的结构性原因，有助于突显AI伦理争论如何发挥作用：它们为AI对系统性生态灭绝的助长作用赋予正当性，使之正常化，而非有意义地挑战它。因

此，我们运用政治生态学视角，加入那些把AI伦理批评为伦理洗白和／或漂绿的论述，主张关注点应转向权力的结构性问题。^[22]

（并不那么）绿色的AI

AI在许多常被忽视的方面都是资源密集型的。近期围绕AI、伦理与可持续性的争论，大多聚焦于训练机器学习模型的碳成本。^[23]这部分源于一篇颇具影响的论文；该论文估计，对某个特定机器学习模型进行一次训练，会排放284,019千克二氧化碳当量（CO₂e）。^[24]虽然这一估算惊人地高，相当于3,422部 iPhone 13 的全生命周期排放量，^[25]但它在原论文中并不突出。原论文承认，替代硬件可使训练时间缩短至原来的1/8.5，表明该估计存在相当大的误差范围。^[26]实际上，后续研究认为，由于采用了错误假设，例如实验使用的是拥有2.13亿参数的“大型”Transformer模型，而非6500万参数的基础模型，该次训练的实际排放量，在平均水平的美国数据中心应为15,200千克CO₂e；如果使用实际模型运行所在地、Google 佐治亚州数据中心的硬件，则为3,200千克CO₂e。^[27]

但实际上，与机器学习系统生态影响更相关的，是 Strubell 等人的另一项发现：他们以案例研究的方式，全程跟踪了另一个模型从开发起始到部署的过程，发现它需要4,789次训练，相当于“60个GPU [图形处理器] 在项目持续的六个月中不间断运转”。^[28]这一发现突显了只关注单次训练二氧化碳排放的谬误。此外，运行机器学习系统——也就是所谓的推理——所用能源，远多于训练模型。Nvidia 估计，机器学习系统成本的80%至90%用于推理，而非初始训练；^[29]Amazon Web Services 则表示，推理占机器学习成本的比例可高达90%。^[30]

数据中心利用高性能计算机并行开展机器学习系统的训练与推理，同时承担数据存储、网站托管及视频转码等活动，因此当然是能源与资源密集使用的场所之一。对全球数据中心能源使用量的估计差异很大，从205太瓦时（TWh）^[31]到400—500太瓦时不等。^[32]这相当于全球全部用电量的1%至2%。数据存储与处理、机器学习数据集规模、训练、推理，以及8K视频转码和流媒体传输等其他计算密集型任务，预计都将增长；因此，到2030年，数据中心能源使用量估计会增加到约780太瓦时。^[33]

过去十年，数据显著向大型云数据中心和超大规模数据中心集中。到2021年，超大规模数据中心已有659个^[34]（Synergy Research Group, 2021），比2015年增加了一倍多。^[35]其中一半仅由三家公司拥有和运营：Amazon、Microsoft 和 Google。^[36]过去许多企业运营小型内部数据中心，而超大规模数据中心则体现了平台资本主义的逻辑：^[37]利用规模经济和网络效应集中设施，由少数寡头科技公司向较小企业出租空间及算力。超大规模数据中心的一项所谓好处，是提高能源效率；^[38]我们将在下一节讨论效率和脱钩时再谈这一点。

碳成本固然重要，但这些研究的狭窄焦点掩盖了一个迫切需要：必须有意义地考虑 AI／机器学习的生态影响，而这些影响远远超出为特定模型训练供电、或在数据中心中进行推理的能源使用。我们需要分析的是，机器学习系统运作所需的整个供应链，以及软件、硬件和基础设施组装体所涉及物质与能源流动。政治生态学在这里为现有“绿色”和“伦理”AI提供了有用的替代进路：它既关注绘制社会技术系统从摇篮到坟墓的物质与能源足迹，也强调生态系统如何与权力、不平等和暴力深刻纠缠并受到它们的影响。数据中心还需要巨量的水。^[39]数据中心内部产生大量热量，必须主动冷却。^[40]水是冷却过程的核心；冷水作为传热机制，被用来降低空气温度。^[41]2014年，美国境内数据中心用水约6260亿升。^[42]尽管数据中心耗水量（每日17亿升）只占总用水量（每日1.218万亿升）的一小部分，^[43]其对可用水资源的急剧、重大影响，却取决于地理位置。例如，美国西部和西南部的数据中心依赖本已缺水、承压的流域，^[44]令科技企业与当地社区及生态系统发生冲突。生态灾难发生时，科技企业往往比当地居民得到更优先的照顾。^[45]

“绿色AI”进路^[46]关注减少数据中心内训练和运行模型的用电，却很少讨论生产或处置硬件、以及容纳并冷却这些技术的基础设施所带来的人类与环境成本。这些社会和环境伤害包括：从刚果民主共和国（DRC）开采钽、钨和黄金等冲突矿产的产业所造成的伤害，^[47]以及依赖童工开采的伤害，例如刚果民主共和国的钴矿开采。^[48]对其他材料，例如镧系元素（通常称为稀土矿物）的无节制开采，会留下影响当地环境与人类健康的有毒遗产，例如在中国造就“癌症村”。^[49]

此外，这些材料从地下采出时，没有一种可以直接用于高性能计算。另一系列能源密集型提取过程必须分离并提纯原料，造成大量额外污染。选矿过程会剔除从地下开采出来的大部分原料。例如，随着高品位铜矿日益枯竭，具有商业开采价值的矿石，其含铜量从2%下降到0.8%，^[50]这意味着，所开采材料质量的99%以上都是废物／尾矿。^[51]提纯过程需要大量能源投入，且往往使用有毒材料。例如，用于CPU、GPU和TPU的硅，必须经过数个化学处理阶段，达到99.999999999%的纯度，^[52]此后再经过热氧化、光刻、等离子蚀刻和掺杂等工序。所有这些工序都大量消耗化学品与能源，必须精密控制物质与温度，其中一些温度高达1100℃。^[53]

AI所需的物质与能源，大多从全球经济边缘地区流向经济核心不断膨胀的技术物质总量（technomass）。^[54]与此同时，数据殖民主义理论描述了金融价值和数据从经济边缘流向核心的过程。^[55]这种生态不平等交换（ecologically unequal exchange, EUE）^[56]并非偶然伴随全球资本主义，而是其构成性要素；它贯穿过去和现在各种形式的殖民主义与帝国主义，持续存在。^[57]在生态不平等交换条件下，原材料和能源从边缘流向核心，而每年产生的5000万吨有毒电子废弃物，又有很大一部分由边缘地区充当其“垃圾场”。^[58]生态经济学近期研究表明，高收入国家的原材料使用量，比本国开采量多出100多亿吨；除高收入国家之外，所有

地区都是原材料的净提供者。^[59]正如我们已经说明的，这也包括AI所需的原材料。

数据中心内的硬件固然是机器学习系统所必需的，但整体性的伦理评估，还应涵盖连接数据中心与终端客户设备的全部网络 and 平台基础设施。这包括海底及陆上光纤电缆、互联网交换点和海缆登陆站，以及大量与机器学习系统交互的台式电脑、笔记本电脑、平板电脑、智能手机及物联网设备。此外，机器学习系统需要训练数据。例如，常用于图像识别的 ImageNet 数据集来自 Flickr，^[60]因此，使用这些数据集的机器学习系统所需的组装体，也包括 Flickr 的平台基础设施，以及创造该数据集的一系列数码相机、智能手机和电脑。图像由通过 Amazon Mechanical Turk 招募的人类标注；这一平台向工人支付的报酬远低于最低工资，同时不提供雇佣关系应有的保障。^[61]与这些系统相关的、受剥削且不稳定的人类劳动，也应被承认为机器学习系统生态的一部分。

描绘当代机器学习所需系统的规模，远远不止于数据中心的一台服务器、它所需的能源及其排放。它包含遍布全球的大量设备、操作、设施和人。每台设备本身也是材料的组装体；这些材料的开采、加工和处置，与一系列环境、社会和劳动正义问题相关，而这些问题在大多数当代“伦理AI”观念中仍遭忽视。承认这个组装体的规模，以及它施加于人类和生态系统的伤害范围与严重程度，就揭穿了下述任何观念：在当前攫取主义资本主义的生态不平等交换安排下，AI是合乎伦理的；或者只要处理隐私问题、消除训练数据集的偏见使其“更公平”，或让模型本身更“透明”、更“可问责”，就能把AI视为合乎伦理。实际上，目前极其狭窄的AI伦理框架，主要有利于科技企业：它们对生态后果灾难性的技术作些微小技术修改，就大胆宣称这足以证明自己的行为合乎伦理。

效率、技术与资本主义

短视地只关注模型训练的能源成本，会产生以提高效率为基础的“解决方案”：通过限制机器学习模型的计算强度或训练数据集规模来降低碳成本。^[62]这些“解决方案”错过了机器学习系统、供应链与基础设施的更广泛问题；只关注数据中心内部的能源使用，既不考虑上下游成本与外部性，也不承认技术、效率与资本主义之间历史形成且持续存在的关系。

“绿色AI”^[63]主张，提高能源效率能够使经济增长与环境影响脱钩，从而减少机器学习造成的生态伤害。评估这一主张，需要了解脱钩、技术与“绿色增长”的一些背景。自工业革命以来，以GDP衡量的经济增长，与温室气体（GHG）排放和物质足迹——即一个社会使用的材料总质量——都存在很强的正相关关系。^[64]脱钩的目标，是在减少当前相关生态伤害的同时，让经济增长继续不减。在这里，必须就脱钩的形式、尺度和速率作出若干关键区分。首先是相对脱钩与绝对脱钩的区别：相对脱钩意味着，虽然增长率与温室气体排放/物质足迹的变化速度出现分

离，资源使用／伤害总量却仍然增加；绝对脱钩则意味着经济增长时，温室气体排放／物质足迹下降。第二项重要区别涉及分析的地理尺度。虽然有许多经合组织（OECD）国家在国家层面实现绝对脱钩的实例，^[65]但这种脱钩是通过产业外移、进口商品与材料实现的，即一种以生态不平等交换为前提的区域脱钩。要避免生态灾难，绝对脱钩必须发生在全球尺度，而非国家或区域尺度。最后，要使绿色增长既合乎伦理又可持续，经济增长与温室气体排放的绝对脱钩必须足够迅速，以达到甚至超越《巴黎协定》等全球环境承诺。

假设经济继续以每年约2%的速度增长，这种脱钩就要求OECD国家以每年15%的速度减少排放，直至归零。^[66]从化石燃料转向可再生能源，确实能使温室气体排放与能源使用脱钩，但没有实证证据支持温室气体排放与经济增长能够以这样的速率实现绝对脱钩。此外，自1990年以来，全球材料使用量的增长超过了GDP的增长。^[67]当代没有任何证据表明，经济增长与资源使用正在趋向脱钩；事实上，在这段时间内，每一单位经济增长都需要更多资源。当前全球物质足迹正接近每年1000亿吨，较1990年增加一倍多，而据估计，可持续水平是每年250亿至500亿吨。^[68]此外，1990年以来物质足迹增长的绝大部分——按人均计算为81%——都归因于高收入国家，进一步表明，对当前生态危机的责任分担极为不平等。^[69]

把目光从全球总量转向效率与AI，数字技术的性能指标与能源使用之间，长期存在相对脱钩。^[70]数据中心体现了这些变化：随着时间推移，尤其是转向超大规模及云数据中心之后，它们的能源效率不断提高。^[71]衡量数据中心能源效率的指标是电能使用效率（power use effectiveness, PUE），该比率表明运行IT设备所需能耗相对于设施总能耗（包括冷却、照明等）的比例。（译注：原文此处的比率表述方向与通常定义相反；PUE通常定义为设施总能耗除以IT设备能耗，这也与下文大于1的数值一致。正文保留原文说法。）理想的PUE为1，表示全部能耗都用于IT设备。到2020年，美国平均PUE为1.59，而某些数据中心已达到1.11。^[72]然而，PUE只衡量效率，并不衡量设施的总能耗，^[73]也不衡量设施总数。因此，尽管设施变得更高效，数据中心的总能耗仍在增长。^[74]

效率提升对于支撑资本主义经济增长至关重要，并贯穿计算发展的全部历史。^[75]问题在于，效率收益通常会被杰文斯悖论等反弹效应抵消：^[76]资源使用效率更高，反而导致资源使用总量上升。不过，提高效率符合数据中心所有者的经济利益，因为这会降低成本，进而增加利润。这并不是说效率“坏”，而是说，在以增长为基础的资本主义经济中，效率带来的节约，通常不会导致排放或材料使用绝对减少——它们会增加总物质投入。正如 Foster 及其同事指出：“一个致力于利润、积累和无止境经济扩张的经济体系，倾向于利用任何效率收益或成本下降来扩大整体生产规模。”^[77]因此，我们需要摆脱把交换价值（即GDP）与社会进步混为一谈的做法，放弃经济无限增长的童话。结论部分将勾勒实现这一转变的策略；

在此之前，我们先转向“绿色资本主义”下解决生态危机的第二种模式：技术解决主义。^[78]

技术解决主义的谬误

新兴技术被定位为治疗生态危机的万能药，让当前的过度消费形式可以照常持续。澳大利亚前总理 Scott Morrison 宣称，我们将通过“技术，而非税收”解决气候变化，便是一个例子。^[79]这一做法远非新鲜事。2007年，在美国退出要求工业化国家减少温室气体排放的《京都议定书》六年之后，时任美国总统 George W. Bush 与时任澳大利亚总理 John Howard 发表联合声明，认为“低排放技术的开发与部署，将是应对气候挑战的关键要素”，^[80]并把清洁煤和碳捕集与封存（CCS）列为技术解决方案。十五年过去了，清洁煤和CCS一次又一次未能证明其技术及经济可行性。^[81]正如 Kuch 所阐明的，CCS的框架来自化石燃料行业的世界观；在这种世界观中，气候变化是一个应该用技术来解决的技术问题，而非需要通过逐步淘汰化石燃料来解决的问题。^[82]

乍看之下，CCS似乎是化石燃料行业与右翼政治领导人用来阻碍应对生态危机行动的障眼法，但政府间气候变化专门委员会（IPCC）第五次评估报告却接受了它。^[83]该报告在旨在避免灾难性气候变化的社会经济模型中，把CCS与生物能源相结合的技术——简称BECCS——置于中心。在116种避免危险升温的情景中，100多种都使用了BECCS。^[84]理论上，BECCS能够产生负排放，因此允许大幅超出将升温控制在2°C以内的碳预算：树木从空气中吸收二氧化碳，树木再被制成颗粒并燃烧，然后把碳排放捕集并储存在地下，从而降低大气CO₂浓度。依赖对21世纪后半叶大量负排放的、存在问题且很可能不准确的预测，使各国政府能够以尚未实现的未来技术创新这一推测性万能药为根据，拖延当下的减排。^[85]虽然CCS技术迄今未能兑现其成效主张，BECCS的生物能源部分也存在问题：IPCC模型中用于生物质的土地，通常需要达到印度面积的一至两倍。^[86]再加上全球变暖导致作物减产，这很可能加剧全球经济边缘部分地区已经遭受的灾难性粮食短缺和生物多样性下降。^[87]

问题不只在于，BECCS等技术解决主义手段不大可能解决生态危机；它们还会转移注意力，积极阻碍迫切需要的集体社会、政治与文化变革，因为它们暗示技术——包括尚不存在的技术——会全面“修复”生态危机，公民现在要求行动也就几乎无益。^[88]就AI而言，所谓解决主义的灵丹妙药，是用可再生能源为数据中心供电。大型科技企业接受了这种极其有限的可持续性定义，并在年度环境报告等宣传材料中突出它，公开把自己描绘成引领社会走向可持续未来的力量。^[89]

例如，Google 的2020年环境报告宣称：“可持续性 is Google 的核心价值之一，自公司二十多年前成立以来，我们一直是气候变化领域的领导者。”^[90]Google 声称，它于2007年成为第一家实现碳中和的大型公司，到2020年已“中和我们自创

立以来遗留的碳足迹，使 Google 成为第一家在整个运营历史中实现碳中和的大型公司”。然而，Google 环境报告末尾的环境数据却呈现出另一幅图景。要解读这些数据，需要对企业衡量和审计排放的方法有基本了解。排放分为范围1、范围2和范围3：^[91]范围1指直接拥有的排放源，例如厂区内锅炉产生的排放；范围2涵盖购入电力及其他公用事业服务所关联的温室气体排放；范围3的排放总量通常最大，包括公司间接负责的全部排放，贯穿其供应链上下游。

Google 报告的2019年排放总量为17,646,902吨CO₂e，其中大部分是范围3排放；文档没有明确其来源，但几乎肯定来自硬件和基础设施的生产。^[92]与此同时，Google 通过现场可再生能源设施、与第三方供应商签订的购电协议，以及碳抵消项目实现的减排总量，为5,725,635吨CO₂e。^[93]简单算术就能表明，净排放量远超抵减量；Google 在2019年并未实现碳中和，更不用说像环境报告宣称的那样抵消公司历史碳足迹。Google 的碳中和宣言只涵盖其微不足道的直接排放与更可观的用电排放，即范围1和范围2，而径直忽略了生产硬件与基础设施所关联的庞大碳足迹。因此，Google 关于碳中和的宣称，是利用选择性披露策略进行漂绿的直白案例：^[94]公司有选择地突出其环境表现的积极部分，从而对自身作出误导性描绘。

与 Google 相比，Apple 把供应链即范围3排放纳入碳足迹，并着手在十年内让全部业务实现碳中和，这一点值得赞许。然而，尽管乍看如此，Apple 距离真正可持续的行动仍有差距。2016年至2020年间，Apple 公司设施——主要是数据中心——的用电量几乎翻倍，从1,420,000兆瓦时增至2,580,000兆瓦时。^[95]虽然这期间 Apple 增加了可再生电力的使用以覆盖增长，但有两个理由说明，不能合理地把它视为可持续。第一，Apple 公司约80%的用电发生在美国，而2021年美国60%以上的发电量来自化石燃料。^[96]脱碳迫切要求用可再生能源替代化石燃料，但在 Apple 的案例中，巨量可再生能源被用于满足数据中心相关用电需求的迅速增长，而非替代现有化石燃料发电。^[97]

能源史常被叙述为煤、石油、天然气依次占据主导的历史。这种分期遮蔽了一个事实：这些能源是彼此补充，而非相互替代。自1990年IPCC首次报告发布以来，煤炭使用量略有增加，石油与天然气使用量几乎翻倍；到2019年，太阳能与风能合计仅提供全球能源的2.6%。^[98]要在避免全球升温超过2度所需的时间内摆脱化石燃料，是一项极其艰巨的任务。这意味着，工业革命以来能源组合不断相加的逻辑，必须让位于快速用可再生能源替代当前化石燃料所供能源的逻辑。在这一背景下，还必须考虑全球边缘及半边缘的非OECD国家增长的能源需求：那里经常停电，数十亿人无法上网。因此，经济核心的富裕国家电力需求迅速增长，既不能被视为公正，也不能被视为“合乎伦理”。

第二个问题是，除了温室气体排放，物质足迹也是迫切的生态问题；太阳能电池板和风力涡轮机需要许多数字硬件同样需要的材料，例如高纯度硅、稀土矿物、锂和

钻。承认这些资源有限，并承认其开采所带来的社会与环境伤害由人们不平等地承受，就意味着必须认识到：太阳与风等能源虽然可以再生，将它们转化为电力的技术手段却不可再生。解决主义者主张，深海采矿^[99]或彗星采矿^[100]等更多推测性技术，将解决绿色资本主义所需陆地材料的稀缺问题。然而，考虑到这些技术涉及数十年的时间跨度，以及深海采矿等活动会造成重大环境伤害的事实，这些主张不应使我们偏离迫切任务：减少排放、减少材料使用，并最终减少消费。

技术解决主义远未真正解决问题；它提供的是极其狭窄的焦点，忽略生态不平等交换的更广泛背景，反而为企业传播内容进行漂绿。新自由主义的技术修补旨在维护既有权力与特权体系：它提出技术修补方法，以维持一种以增长为导向的资本主义模式，而这一模式从根本上不合乎伦理，也与生态正义相悖。

结论

本章对一种主张提出了严厉批评：认为“伦理AI”和“绿色AI”能够使当前AI系统变得合乎伦理且公正，尤其是在伦理干预主要局限于能够用数学处理的技术修补、旨在让AI系统更加“公平”、“可问责”和“透明”的情况下。^[101]相反，我们必须采取行动，补救AI相关的生态伤害，而这些行动要处理AI深深卷入其中的、极不公正、不可持续且不平等的社会经济体系。^[102]根本而言，按照政治生态学的规范性社会伦理模型，只要AI仍与生态灭绝共谋，就不能被视为合乎伦理。因此，结尾我们勾勒一些新出现的后资本主义进路，它们重新想象并预先实践破坏生态更少、更“合乎伦理”的社会技术系统。

资本主义经济的结构，要求GDP每年至少复合增长2%至3%。^[103]经济增长与物质足迹、温室气体排放高度相关。可再生能源的广泛采用虽然能使温室气体排放与GDP脱钩，但能源与GDP之间的基础相关性依然存在，而“可再生”能源仍需要大量不可再生材料。这意味着，在物质有限的地球上，资本主义生产方式必然追逐的无限增长不可能实现。因此，后资本主义立场普遍赞同：有必要以其他指标取代GDP，衡量社会经济福祉。^[104]GDP衡量经济交换价值，即民族国家内部货币交易的总和；它广受批评，因为它把燃煤发电站、生产核武器和住院人数增加等破坏性的事物计为价值。与此同时，GDP不为那些不产生或不包含交换价值的事物计价，包括清洁空气、生物多样性、无偿家务劳动，以及维基百科等以共有资源为基础的数字项目。同样，GDP也不反映民族国家内部惊人的不平等程度。因此，以GDP为社会进步指标，就无法认识到：自20世纪70年代以来，发达经济体的增长收益几乎全归最富裕者所有，而贫困水平一直稳步上升。^[105]所以，后资本主义论述明确指出，以交换价值为中心的GDP，不足以衡量生活中真正重要的东西。

政治生态学内部，对去增长、生态现代主义与生态社会主义进路的优点及潜在缺陷已有大量争论；^[106]不过，生态社会主义者^[107]与去增长倡导者^[108]都主张让人类与生态系统之间的关系去商品化，并强调需要以使用价值为基础的经济。然而，讨

论技术时，不同立场之间存在张力：生态社会主义立场主张民主化的集中化，例如国家资助的全国性数字公共服务和基础设施，^[109]或平台社会主义；^[110]去增长进路则强调共生性，把它视为创造非异化技术的方法，借助各种共用实践形式，^[111]包括自由／开源软件与点对点运动。^[112]我们无意倡导某种特定“解决方案”，而是简要突显后资本主义进路的多样性，以推进那些致力于推动各种以生态系统为中心的替代方案者之间的对话与团结，替代技术解决主义和超高效率的“绿色”资本主义。

就目前而言，AI和数据中心表现出网络效应与规模经济，导向超大规模数据中心所体现的极度集中；这反过来又造成企业对这些系统和基础设施的寡头控制。^[113]虽然AI、互联网与计算机都在资本主义经济中发展出来，但目前企业主导的AI基础设施模式并非不可避免。实际上，后资本主义进路的一项关键主张，就是改变技术治理能够改变技术运作方式，从而显著减少对生态系统的伤害和社会不平等。平台社会主义者认为，当前企业与生态系统——包括人类——之间的权力不对称，可以通过集中化的公共服务模式得到有意义的处理，例如许多国家用于医疗、通信、污水系统及其他基础设施的模式。当“自然垄断”^[114]出现时，以使用价值为导向的进路主张把基础设施作为社会化的公共财产来管理，而非作为私人商品。另一些进路则通过合作社、联邦式和分布式点对点系统，以及其他形式的数字共有资源，提出去中心化的替代方案。此外，不再以交换价值衡量财富，意味着许多有害的AI形式将不再被视为有用或有价值。

消除有害的资本主义监控模式——这些模式旨在助推公民进行不可持续的消费——以及企业尽可能囤积数据、寄望未来找到变现办法的做法，^[115]将使数据存储与计算处理量能够显著下降。这有助于减少AI当前造成的社会及环境伤害，同时为某些领域的增长提供空间：在这些领域，AI可以支持生态系统与社区的修复，而不是让掌控技术及相关基础设施者的私人利益受惠。在数据／数字殖民主义背景下，^[116]这一变革的关键部分必须包括数据去殖民化^[117]和增强技术主权，^[118]使个人与社区能够利用自己的数据谋求自身利益。不过，去殖民化也必须超越数据，处理生态不平等交换。

必须突显AI在当前攫取主义资本主义体系中的作用：这一体系以全球各地区内部及地区之间惊人的社会不平等为基础，通过贪婪的过度消费，造成全球生态崩溃；它既不能被视为公正，也不能被视为合乎伦理。在处理气候与生态债务的同时，迫切需要让经济核心的经济体系迅速去商品化。这将使生活在核心以外的人们能够重新确立对其资源、技术、能源和土地的主权。虽然对于以何种适当方式超越资本主义，存在充满争议的辩论，但起点是承认：我们能够做到。政治生态学视角要求我们加深理解AI在后资本主义转型过程中可能发挥的作用，并围绕“AI伦理”与正义发展一种纳入这些广泛目标的实践。面对全球性的缺陷结构与系统失灵，我们需要的远不止一块遮盖问题的伦理膏药。

资助披露

作者未就本文获得外部资助，也没有需要披露的财务利益冲突。

参考文献

- ABC News, ‘PM defends his “technology not taxes” approach to cutting emissions’, *ABC News*, 23 April 2021, <https://www.abc.net.au/news/2021-04-23/pm-defends-his-technology-not-taxes-approach-to-emissions/13315616>. [中文题名：总理为其“技术而非税收”的减排路径辩护。]
- Al Qundus, Jamal, Dabbour, Kossai, Gupta, Shivam, Meissonier, Régis, and Paschke, Adrian. ‘Wireless Sensor Network for AI-based Flood Disaster Detection’, *Annals of Operations Research*, 7 August 2020: 1–23. [中文题名：用于基于AI的洪水灾害检测的无线传感器网络。]
- Amoore, Louise. *Cloud Ethics: Algorithms and the Attributes of Ourselves and Others*, Durham: Duke University Press, 2020. [中文题名：云伦理：算法与我们自身及他人的属性。]
- Anderson, Kevin and Peters, Glen. ‘The Trouble with Negative Emissions’, *Science* 354. 6309 (2016): 182–3. [中文题名：负排放的问题。]
- Andrae, Anders, S.E. ‘Hypotheses for Primary Energy Use, Electricity Use and CO₂ Emissions of Global Computing and its Shares of the Total Between 2020 and 2030’, *WSEAS Transactions on Power Systems* 15 (2020): 50–9. [中文题名：关于2020至2030年全球计算的一次能源使用、电力使用和CO₂排放及其占总量比例的假设。]
- Andrejevic, Mark. *Automated Media*, New York: Routledge, 2020. [中文题名：自动化媒体。]
- Apple. *iPhone 13 Product Environmental Report*, 14 September 2021, https://www.apple.com/environment/pdf/products/iphone/iPhone_13_PER_Sept2021.pdf. [中文题名：iPhone 13产品环境报告。]
- . *Environmental Progress Report*, April 2021, https://www.apple.com/environment/pdf/Apple_Environmental_Progress_Report_2021.pdf. [中文题名：环境进展报告。]
- Barabási, Albert-László and Albert, Réka. ‘Emergence of Scaling in Random Networks’, *Science* 286.5439 (1999): 509–12. [中文题名：随机网络中标度的涌现。]
- Barbrook, Richard and Cameron, Andy. ‘The Californian Ideology’, *Science as Culture* 6.1 (1996): 44–72. [中文题名：加利福尼亚意识形态。]
- Barnes, Elizabeth A., Hurrell, James W., Ebert-Uphoff, Imme., Anderson, Chuck and Anderson, David. ‘Viewing Forced Climate Patterns Through an AI Lens’, *Geophysical Research Letters*, 46.22 (2019): 13389–98. [中文题名：透过AI视角观察受迫气候模式。]
- Barr, Jeff. *Amazon EC2 Update—Inf1 Instances with AWS Inferentia Chips for High Performance Cost-Effective Inferencing*, 3 December 2019, <https://aws.amazon.com/blogs/aws/amazon-ec2-update-inf1-instances-with-aws-inferentia-chips-for-high->

- performance-cost-effective-inferencing/. [中文题名: Amazon EC2更新——配备AWS Inferentia芯片、实现高性能与高成本效益推理的Inf1实例。]
- Bashroush, Rabih and Lawrence, Andy. *Beyond PUE: Tackling Its Wasted Terawatts*, Uptime Institute, January 2020, <https://uptimeinstitute.com/beyond-pue-tackling-wasted-terawatts>. [中文题名: 超越PUE: 应对其浪费的太瓦级电力。]
- Bastani, Aaron. *Fully Automated Luxury Communism*, London: Verso Books, 2020. [中文题名: 全自动奢华共产主义。]
- Bedford, Laura, Mann, Monique, Walters, Reece, and Foth, Marcus. 'A Post-capitalocentric Critique of Digital Technology and Environmental Harm: New Directions at the Intersection of Digital and Green Criminology, in *Beyond Cybercrime: New perspectives on crime, harm and digital technologies*, *International Journal for Crime, Justice and Social Democracy (Special Issue)*, 11.1 (2022): 167–81. [中文题名: 对数字技术与环境伤害的后资本中心主义批判: 数字犯罪学与绿色犯罪学交叉领域的新方向; 载《超越网络犯罪: 犯罪、伤害与数字技术的新视角》。]
- Beer, David. *The Data Gaze: Capitalism, Power and Perception*, Newbury Park: Sage, 2018. [中文题名: 数据凝视: 资本主义、权力与感知。]
- Beinstein, Andreas. 'Conviviality, the Internet, and AI. Ivan Illich, Bernard Stiegler, and the Question Concerning Information-technological Self-limitation', *Open Cultural Studies* 4.1 (2020): 131–42. [中文题名: 共生性、互联网与AI: Ivan Illich、Bernard Stiegler以及信息技术自我限制问题。]
- Beiser, Vince. *The World in a Grain: The Story of Sand and How It Transformed Civilization*. New York: Riverhead Books, 2018. [中文题名: 一粒沙中的世界: 沙子的故事及其如何改变文明。]
- Bender E.M., Gebu, Timnit, McMillan-Major, Angelina, and Shmitchell, Shmargaret. 'On the Dangers of Stochastic Parrots: Can Language Models Be Too Big??' *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 1 March 2021, 610–23. [中文题名: 论随机鹦鹉的危险: 语言模型是否可能过于庞大? ?。]
- Benkler, Yochai. 'Don't Let Industry Write the Rules for AI', *Nature* 569.7754 (2019): 161–2. [中文题名: 别让行业制定AI规则。]
- Bergvall-Kåreborn, Birgitta and Howcroft, Debra. 'Amazon Mechanical Turk and the Commodification of Labour', *New Technology, Work and Employment* 29.3 (2014): 213–23. [中文题名: Amazon Mechanical Turk与劳动商品化。]
- Bonneuil, Christophe and Fressoz, Jean-Baptiste. *The Shock of the Anthropocene: The Earth, History and Us*, London: Verso Books, 2016. [中文题名: 人类世的冲击: 地球、历史与我们。]
- Brady, Gemma A., Kapur, Nikil, Summers, Jonathan L., and Thompson, Harvey M. 'A Case Study and Critical Assessment in Calculating Power Usage Effectiveness for a Data Centre', *Energy Conversion and Management* 76 (2013): 155–61. [中文题名: 数据中心电能使用效率计算的案例研究与批判性评估。]

- Bridle, James. *New Dark Age: Technology and the End of the Future*, London: Verso Books, 2018. [中文题名: 新黑暗时代: 技术与未来的终结。]
- Bringezu, Stefan. 'Possible Target Corridor for Sustainable Use of Global Material Resources', *Resources* 4.1 (2015): 25–54. [中文题名: 全球物质资源可持续使用的可能目标区间。]
- Bryant, Raymond L. and Jarosz, Lucy. 'Ethics in Political Ecology: A Special Issue of Political Geography: Introduction: Thinking About Ethics in Political Ecology', *Political Geography* 23. 7 (2004): 807–12. [中文题名: 政治生态学中的伦理: 《政治地理学》专刊导论——思考政治生态学中的伦理。]
- Bunker, Stephan G. 'Raw Materials and the Global Economy', *Society and Natural Resources* 9.4 (1996): 419–29. [中文题名: 原材料与全球经济。]
- Bush, George W. and Howard, John. *Joint Statement on Climate Change and Energy by President Bush and Australian Prime Minister John Howard*, 4 September 2007, <https://georgewbush-whitehouse.archives.gov/news/releases/2007/09/20070904-8.html>. [中文题名: 布什总统与澳大利亚总理John Howard关于气候变化和能源的联合声明。]
- Chang, Li-Chiu, Fi-John Chang, Shun-Nien Yang, Fong-He Tsai, Ting-Hua Chang, and Edwin E. Herricks. 'Self-Organizing Maps of Typhoon Tracks Allow for Flood Forecasts up to Two Days in Advance', *Nature Communications* 11.1 (2020): 1983. [中文题名: 台风路径的自组织映射使提前两天预报洪水成为可能。]
- Couldry, Nick, and Mejias, Ulises A. *The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism*, Stanford: Stanford University Press, 2019. [中文题名: 连接的代价: 数据如何殖民人类生活并将其据为资本主义所有。]
- . 'Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject', *Television & New Media* 20.4 (2018): 336–49. [中文题名: 数据殖民主义: 重新思考大数据与当代主体的关系。]
- Crawford, Kate. 'The Atlas of AI: Power Politics and the Planetary Costs of Artificial Intelligence', New Haven: Yale University Press, 2021. [中文题名: AI地图集: 权力政治与人工智能的行星代价。]
- Daly Angela, Hagendorff, Thilo, Hui, Li, Mann, Monique, Marda, Vidushi, Wagner, Ben, Wei Wang, Wayne. 'AI, Governance and Ethics: Global Perspectives' in Pollicino, Oreste and de Gregorio, Giovanni (eds) *Constitutional Challenges in the Algorithmic Society*, Cambridge: Cambridge University Press, 2021. [中文题名: AI、治理与伦理: 全球视角; 载《算法社会中的宪制挑战》。]
- Daly, Angela, Devitt, S, Kate, and Mann, Monique. 'AI Ethics Needs Good Data' in Pieter Verdegem (ed) *AI for Everyone? Critical Perspectives*, London: University of Westminster Press, 2021. [中文题名: AI伦理需要良好数据; 载《人人共享AI? 批判性视角》。]
- de Freitas Netto, Sebastião Vieira, Sobral, Marcos Felipe Falcão, Ribeiro, Ana Regina Bezerra, and Soares, Gleibson Robert da Luz. 'Concepts and Forms of Greenwashing: A Systematic Review', *Environmental Sciences Europe* 32.1 (2020): 1–12. [中文题名: 漂绿的概念与形式: 系统综述。]

- Delmas, Magali A., and Cuerel Burbano, Vanessa. 'The Drivers of Greenwashing', *California Management Review* 54.1 (2011): 64–87. [中文题名：漂绿的驱动力。]
- Dieter, Cheryl A., Maupin, Molly A., Caldwell, Rodney R., Harris, Melissa A., Ivahnenko, Tamara I., Lovelace, John K., Barber, Nancy L., and Linsey, Kristin S. *Estimated Use of Water in the United States in 2015*, Reston: U.S. Geological Survey, 2018. [中文题名：2015年美国用水量估算。]
- Dorninger, Christian, Hornborg, Alf, Abson, David J., Von Wehrden, Henrik, Schaffartzik, Anke, Giljum, Stefan, Engler, John-Oliver, et al. 'Global Patterns of Ecologically Unequal Exchange: Implications for Sustainability in the 21st Century', *Ecological Economics* 179 (2021): 106824. [中文题名：生态不平等交换的全球模式：对21世纪可持续性的启示。]
- 'The World's Most Valuable Resource Is No Longer Oil, but Data', *The Economist*, 6 May 2017, <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>. [中文题名：世界上最有价值的资源不再是石油，而是数据。]
- Eubanks, Virginia. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*, New York: St. Martin's Press, 2018. [中文题名：自动化不平等：高科技工具如何给穷人画像、管控和惩罚穷人。]
- Forti, Vanessa, Balde, Cornelis P., Kuehr, Ruediger, and Bel, Garam. *The Global E-Waste Monitor 2020: Quantities, Flows and the Circular Economy Potential*, Bonn, Geneva and Rotterdam: United Nations University/United Nations Institute for Training and Research, International Telecommunication Union, and International Solid Waste Association, 2020. [中文题名：2020年全球电子废弃物监测：数量、流动与循环经济潜力。]
- Foster, John Bellamy, Clark, Brett, and York, Richard. 'Capitalism and the Curse of Energy Efficiency', *Monthly Review* 62.6 (2010): 1–12. [中文题名：资本主义与能源效率的诅咒。]
- Friggens, Nina L., Hester, Alison J., Mitchell, Ruth J., Parker, Thomas C., Subke, Jens-Arne, and Wookey, Philip A. 'Tree Planting in Organic Soils Does Not Result in Net Carbon Sequestration on Decadal Timescales', *Global Change Biology* 26.9 (2020): 5178–88. [中文题名：在有机土壤中植树不会在十年尺度上产生净碳封存。]
- Garnham, Nicholas. 'Editorial', *Media, Culture & Society* 14.3 (1992): 339–42. [中文题名：社论。]
- Gómez-Baggethun, Erik. 'More Is More: Scaling Political Ecology within Limits to Growth', *Political Geography* 76 (2020): 102095. [中文题名：更多就是更多：在增长极限内拓展政治生态学的尺度。]
- Google. 'Google Environmental Report 2020', 2020, <https://www.gstatic.com/gumdrop/sustainability/google-2019-environmental-report.pdf>. [中文题名：Google 2020年环境报告。]
- Hagendorff, Thilo. 'The Ethics of AI Ethics: An Evaluation of Guidelines', *Minds & Machines* 30 (2020): 99–12. [中文题名：AI伦理的伦理：指南评估。]

- Harvey, David. *The Enigma of Capital and the Crises of Capitalism*. London: Profile Books, 2010. [中文题名: 资本之谜与资本主义危机。]
- Hickel, Jason. *Less Is More: How Degrowth Will Save the World*, New York: Penguin Random House, 2020. [中文题名: 少即是多: 去增长将如何拯救世界。]
- Hickel, Jason, and Kallis, Giorgos. 'Is Green Growth Possible?' , *New Political Economy* (2019): 1–18. [中文题名: 绿色增长可能吗? 。]
- Higgins, Polly. *Eradicating Ecocide: Exposing the Corporate and Political Practices Destroying the Planet and Proposing the Laws to Eradicate Ecocide*, London: Shepard-Walwyn, 2016. [中文题名: 根除生态灭绝: 揭露摧毁地球的企业与政治实践, 并提出根除生态灭绝的法律。]
- Hogan, Mél. 'Data Flows and Water Woes: The Utah Data Center' , *Big Data & Society* 2 (2): (2015): 1–12. [中文题名: 数据流与水之忧: 犹他数据中心。]
- . 'Big Data Ecologies' , *Ephemeria* 18(3) (2018): 631–57. [中文题名: 大数据生态。]
- Hogan, Mél, and Asta Vonderau. 'The Nature of Data Centers' , *Culture Machine* 18 (2019) 1–4. [中文题名: 数据中心的自然。]
- Holl, Karen D., and Brancalion, Pedro H. S. 'Tree Planting Is Not a Simple Solution' , *Science* 368.6491 (2020): 580–1. [中文题名: 植树并非简单解决方案。]
- Hornborg, Alf. 'Towards an Ecological Theory of Unequal Exchange: Articulating World System Theory and Ecological Economics' , *Ecological Economics* 25.1 (1998): 127–36. [中文题名: 走向不平等交换的生态理论: 联结世界体系理论与生态经济学。]
- . *The Power of the Machine: Global Inequalities of Economy, Technology, and Environment*, Lanham: Altamira Press, 2001. [中文题名: 机器的力量: 经济、技术与环境的全球不平等。]
- . 'The Political Ecology of the Technocene' , *The Anthropocene and the Global Environmental Crisis: Rethinking Modernity in a New Epoch*, (2015): 177–83. [中文题名: 技术世的政治生态学; 载《人类世与全球环境危机: 在新纪元重新思考现代性》。]
- . 'Energy, Space, and Movement: Toward a Framework for Theorizing Energy Justice' , *Geografiska Annaler: Series B, Human Geography* 102.1 (2020): 8–20. [中文题名: 能源、空间与运动: 迈向能源正义理论化的框架。]
- Hornborg, Alf and Martinez-Alier, Joan. 'Ecologically Unequal Exchange and Ecological Debt' , *Journal of Political Ecology* 23.1(2016): 328–33. [中文题名: 生态不平等交换与生态债务。]
- Huntingford, Chris, Jeffers, Elizabeth S., Bonsall, Michael B., Christensen, Hannah M., Lees, Thomas, and Yang, Hui. 'Machine Learning and Artificial Intelligence to Aid Climate Change Research and Preparedness' , *Environmental Research Letters* 14.12 (2019): 124007. [中文题名: 帮助气候变化研究及准备工作的机器学习与人工智能。]
- Jackson, Tim. *Post Growth: Life After Capitalism*, Cambridge: Polity, 2020. [中文题名: 后增长: 资本主义之后的生活。]

- Jevons, William Stanley. (1865) *The Coal Question*, 3rd edition. New York: Augustus M. Kelley, 1965. [中文题名: 煤炭问题。]
- Jiao, Pengcheng, and Alavi, Amir H. ‘Artificial Intelligence in Seismology: Advent, Performance and Future Trends’, *Geoscience Frontiers* 11.3 (2020): 739–44. [中文题名: 地震学中的人工智能: 出现、表现与未来趋势。]
- Kallis, Giorgos. *Degrowth*, Newcastle-Upon-Tyne: Agenda Publishing, 2018. [中文题名: 去增长。]
- Keane, Pearse A., and Topol, Eric J. ‘With an Eye to AI and Autonomous Diagnosis’, *Npj Digital Medicine* 1.1 (2018): 40. [中文题名: 着眼于AI与自主诊断。]
- Kerschner, Christian, Wächter, Petra, Nierling, Linda, and Ehlers, Melf-Hinrich, ‘Degrowth and Technology: Towards Feasible, Viable, Appropriate and Convivial Imaginaries’, *Journal of Cleaner Production*, 197 (2018): 1619–36. [中文题名: 去增长与技术: 迈向可行、可存续、適切且共生的想象。]
- Klinger, Julie Michelle. *Rare Earth Frontiers*, Ithaca, London: Cornell University Press, 2017. [中文题名: 稀土边疆。]
- Koppelaar, Rembrandt H. E. M., and Koppelaar, Hendrik. ‘The Ore Grade and Depth Influence on Copper Energy Inputs’, *BioPhysical Economics and Resource Quality* 1.2 (2016): 1–16. [中文题名: 矿石品位与埋深对铜生产能源投入的影响。]
- Kostakis, Vasilis, Latoufis, Kostas, Liarakapis, Minas, and Bauwens, Michel. ‘The Convergence of Digital Commons with Local Manufacturing from a Degrowth Perspective: Two Illustrative Cases’, *Journal of Cleaner Production* 197 (2018): 1684–93. [中文题名: 从去增长视角看数字共有资源与本地制造的融合: 两个说明性案例。]
- Kostakis, Vasilis, Roos, Andreas and Bauwens, Michel. ‘Towards a Political Ecology of the Digital Economy: Socio-Environmental Implications of Two Competing Value Models’, *Environmental Innovation and Societal Transitions* 18 (2016): 82–100. [中文题名: 走向数字经济的政治生态学: 两种竞争价值模式的社会环境含义。]
- Kovel, Joel. *The Enemy of Nature: The End of Capitalism or the End of the World?*, London: Zed Books, 2007. [中文题名: 自然之敌: 资本主义的终结还是世界的终结?。]
- Kuch, Declan. ‘“Fixing” Climate Change through Carbon Capture and Storage: Situating Industrial Risk Cultures’, *Futures* 92 (2017): 90–99. [中文题名: 通过碳捕集与封存“修复”气候变化: 定位工业风险文化。]
- Kukutai, Tahu, and Taylor, John (eds). *Indigenous Data Sovereignty: Toward an Agenda*, Canberra: ANU press, 2016. [中文题名: 原住民数据主权: 迈向一项议程。]
- Lacoste, Alexandre, Luccioni, Alexandra, Schmidt, Victor, and Dandres, Thomas. ‘Quantifying the Carbon Emissions of Machine Learning.’ arXiv preprint arXiv: 1910.09700, (2019). [中文题名: 量化机器学习的碳排放。]
- Leopold, George. ‘Aws to Offer Nvidia’ s T4 Gpus for Ai Inferencing’, *HPC Wire*, 19 March 2019, <https://www.hpcwire.com/2019/03/19/aws-upgrades-its-gpu-backed-ai-inference-platform/>. [中文题名: AWS将提供Nvidia的T4 GPU用于AI推理。]

- Malm, Andreas, and Hornborg, Alf. 'The Geology of Mankind? A Critique of the Anthropocene Narrative', *The Anthropocene Review* 1.1 (2014): 62–9. [中文题名：人类的地质学？对人类世叙事的批判。]
- Mann, Monique and Daly, Angela. '(Big) Data and the North-in-South: Australia's Informational Imperialism and Digital Colonialism', *Television & New Media* 20.4 (2019): 379–95. [中文题名：（大）数据与南方中的北方：澳大利亚的信息帝国主义和数字殖民主义。]
- Mann, Monique, Mitchell, Peta, Foth, Marcus, and Anastasiu, Irina. '#BlockSidewalk to Barcelona: Technological Sovereignty and the Social Licence to Operate Smart Cities', *Journal of the Association for Information Science and Technology* 71.9 (2020): 1103–15. [中文题名：从#BlockSidewalk到巴塞罗那：技术主权与智慧城市运营的社会许可。]
- Masanet, Eric, Shehabi, Arman, Lei, Nuoa, Smith, Sarah, and Koomey, Jonathan. 'Recalibrating Global Data Center Energy-Use Estimates', *Science* 367.6481 (2020): 984–6. [中文题名：重新校准全球数据中心能耗估算。]
- Massumi, Brian. *99 Theses on the Revaluation of Value: A Postcapitalist Manifesto*, Minneapolis: University of Minnesota Press, 2018. [中文题名：关于价值重估的99条论纲：后资本主义宣言。]
- Mathai, Manu V., Isenhour, Cindy, Stevis, Dimitris, Vergragt, Philip, Bengtsson, Magnus, Lorek, Sylvia, Fogh Mortensen, Lars, et al. 'The Political Economy of (un) Sustainable Production and Consumption: A Multidisciplinary Synthesis for Research and Action', *Resources, Conservation and Recycling* 167 (2021): 105265. [中文题名：（不）可持续生产与消费的政治经济学：面向研究与行动的多学科综合。]
- Moore, Jason W. *Anthropocene or Capitalocene?: Nature, History, and the Crisis of Capitalism*. Oakland: Pm Press, 2016. [中文题名：人类世还是资本世？自然、历史与资本主义危机。]
- Morozov, Evgeny. *To Save Everything, Click Here: The Folly of Technological Solutionism*, London: Penguin, 2014. [中文题名：要拯救一切，请点击这里：技术解决主义的愚妄。]
- Muldoon, James. 'Platform Socialism': *How to Reclaim our Digital Future from Big Tech*, London: Pluto Press, 2022. [中文题名：平台社会主义：如何从大型科技企业手中夺回我们的数字未来。]
- Mytton, David. 'Data Centre Water Consumption', *Npj Clean Water* 4.1 (2021): 1–6. [中文题名：数据中心用水量。]
- Neilson, Brett, and Rossiter, Ned. 'Automating Labour and the Spatial Politics of Data Centre Technologies', in Mascha Will-Zocholl and Caroline Roth-Ebner (eds) *Topologies of Digital Work*, London: Palgrave Macmillan, Cham, 2021, pp. 77–101. [中文题名：劳动自动化与数据中心技术的空间政治；载《数字工作的拓扑》。]
- Niarchos, Nicholas. 'The Dark Side of Congo's Cobalt Rush', *The New Yorker*, 2021. <https://www.newyorker.com/magazine/2021/05/31/the-dark-side-of-congos-cobalt-rush>. [中文题名：刚果钴矿热潮的阴暗面。]

- Noble, Safiya Umoja. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: NYU Press, 2018. [中文题名: 压迫的算法: 搜索引擎如何强化种族主义。]
- Nost, Eric, and Goldstein, Jenny Elaine. 'A Political Ecology of Data', *Environment and Planning E: Nature and Space* (2021): 25148486211043503. [中文题名: 数据的政治生态学。]
- Nost, Eric, and Colven, Emma. 'Earth for AI: A Political Ecology of Data-Driven Climate Initiatives', *Geoforum* 130 (2022): 23–34. [中文题名: 为AI而存在的地球: 数据驱动气候行动的政治生态学。]
- Patel, Raj, and Moore, Jason W. *A History of the World in Seven Cheap Things: A Guide to Capitalism, Nature, and the Future of the Planet*, New York: Verso, 2017. [中文题名: 以七种廉价之物书写的世界史: 资本主义、自然与地球未来指南。]
- Patterson, David, Gonzalez, Joseph, Le, Quoc, Liang, Chen, Munguia, Lluís-Miquel, Rothchild, Daniel, So, David, Texier, Maud, and Dean, Jeff. 'Carbon Emissions and Large Neural Network Training', arXiv preprint arXiv:2104.10350, 2021. [中文题名: 碳排放与大型神经网络训练。]
- Phan, Thao, and Wark, Scott. 'What Personalisation Can Do for You! Or: How to do Racial Discrimination without "Race".' *Culture machine* 20 (2021): 1–29. [中文题名: 个性化能为你做什么! 或者: 如何在没有“种族”的情况下进行种族歧视。]
- Pickren, Graham. 'Geographies of E - Waste: Towards a Political Ecology Approach to E - Waste and Digital Technologies', *Geography Compass* 8.2 (2014): 111–24. [中文题名: 电子废弃物地理学: 走向电子废弃物与数字技术的政治生态学进路。]
- Piketty, Thomas. *Capital in the Twenty-first Century*, Cambridge: Harvard University Press, 2018. [中文题名: 21世纪资本论。]
- Ranganathan, Janet, Corbier, Laurent, Bhatia, P., Schmitz, Simon, Gage, Peter, and Oren, Kjell. 'The Greenhouse Gas Protocol: A Corporate Accounting and Reporting Standard (revised edition)', Washington, DC: World Resources Institute and World Business Council for Sustainable Development, 2004. [中文题名: 温室气体核算体系: 企业核算与报告标准 (修订版)。]
- Riemens, Rianne. 'Decoupling as Rhetorical Strategy in Google's Green Discourse', *AoIR Selected Papers of Internet Research*, 2021. [中文题名: 作为Google绿色话语修辞策略的脱钩。]
- Ritchie, Hannah and Roser, Max. *Energy Mix*, 2021, <https://ourworldindata.org/energy-mix>. [中文题名: 能源组合。]
- Robbins, Paul. 'Is Less More...or Is More Less? Scaling the Political Ecologies of the Future', *Political Geography* 76 (2020): 102018. [中文题名: 少即是多……还是多即是少? 未来政治生态学的尺度拓展。]
- Sadowski, Jathan. 'When Data Is Capital: Datafication, Accumulation, and Extraction', *Big Data & Society* 6.1 (2019): 1–12. [中文题名: 当数据成为资本: 数据化、积累与攫取。]

- . ‘The Internet of Landlords: Digital Platforms and New Mechanisms of Rentier Capitalism’, *Antipode* 52 (2020): 562–80. [中文题名：房东的互联网：数字平台与食利资本主义的新机制。]
- Sadowski, Jathan, Viljoen, Salomé, and Whittaker, Meredith. ‘Everyone Should Decide How their Digital Data are Used—Not Just Tech Companies’, *Nature* 595 (2021): 169–171. [中文题名：人人都应决定自己的数字数据如何使用，而不仅仅是科技公司。]
- Sanyé-Mengual, Esther, Secchi, Michaela, Corrado, Sara, Beylot, Antoine, and Sala, Serenella. ‘Assessing the Decoupling of Economic Growth from Environmental Impacts in the European Union: A Consumption-based Approach’, *Journal of Cleaner Production* 236 (2019): 117535. [中文题名：评估欧盟经济增长与环境影响的脱钩：基于消费的进路。]
- Schwartz, Roy, Dodge, Jesse, Smith, Noah A., and Etzioni, Oren. ‘Green AI’, *Communications of the ACM* 63.12 (2020): 54–63. [中文题名：绿色AI。]
- Seddon, Nathalie, Smith, Alison, Smith, Pete, Key, Isabel, Chausson, Alexandre, Girardin, Cécile, House, Jo, Srivastava, Shilpi, and Turner, Beth. ‘Getting the Message Right on Nature-Based Solutions to Climate Change’, *Global Change Biology* 27.8 (2021): 1518–46. [中文题名：准确传达以自然为基础的气候变化解决方案的信息。]
- Shehabi, Arman, Smith, Sarah J., Masanet, Eric, and Koomey, Jonathan. ‘Data Center Growth in the United States: Decoupling the Demand for Services from Electricity Use’, *Environmental Research Letters* 13.12 (2018): 124030. [中文题名：美国数据中心增长：服务需求与用电脱钩。]
- Shehabi, Arman, Smith, Sarah, Sartor, Dale, Brown, Richard, Herrlin, Magnus, Koomey, Jonathan, Masanet, Eric, Horner, Nathaniel, Azevedo, Inês, and Lintner, William. *United States Data Center Energy Usage Report*, U.S. Department of Energy Office of Scientific and Technical Information, 2016, <https://www.osti.gov/biblio/1372902/>. [中文题名：美国数据中心能源使用报告。]
- Shell. *Artificial Intelligence*, 2021, <https://www.shell.com/energy-and-innovation/digitalisation/digital-technologies/shell-ai.html>. [中文题名：人工智能。]
- Siddik, Md Abu Bakar, Shehabi, Arman, and Marston, Landon. ‘The Environmental Footprint of Data Centers in the United States’, *Environmental Research Letters* 16.6 (2021): 064017. [中文题名：美国数据中心的环境足迹。]
- Sluis Katarina, and Li, Fei-Fei. *Where Did ImageNet Come From?*, 2019, <https://unthinking.photography/articles/where-did-imagenet-come-from>. [中文题名：ImageNet从何而来?。]
- Srnicek, Nick. *Platform Capitalism*, Cambridge: Polity, 2016. [中文题名：平台资本主义。]
- Starosielski, Nicole. ‘Thermocultures of Geological Media’, *Cultural Politics* 12.3 (2016): 293–309. [中文题名：地质媒介的热文化。]
- Stiegler, Bernard. *Automatic Society: The Future of Work*, Cambridge: Polity, 2017. [中文题名：自动化社会：工作的未来。]

- Strubell, Emma, Ananya Ganesh, and Andrew McCallum. 'Energy and Policy Considerations for Deep Learning in Nlp', arXiv preprint arXiv:1906.02243, 2019. [中文题名: 自然语言处理深度学习的能源与政策考量。]
- Swinhoe, Dan. 'Microsoft, Amazon, and Google Operate Half the World's 600 Hyperscale Data Centers,' , 2021, <https://www.datacenterdynamics.com/en/news/microsoft-amazon-and-google-operate-half-the-worlds-600-hyperscale-data-centers/>. [中文题名: Microsoft、Amazon与Google运营着全球600个超大规模数据中心中的一半。]
- Taffel, Sy. 'Hopeful Extinctions? Tesla, Technological Solutionism and the Anthropocene', *Culture Unbound: Journal of Current Cultural Research* 10.2 (2018): 163–84. [中文题名: 充满希望的灭绝? Tesla、技术解决主义与人类世。]
- . *Digital Media Ecologies*, New York and London: Bloomsbury, 2019. [中文题名: 数字媒介生态。]
- . 'Data and Oil: Metaphor, Materiality and Metabolic Rifts', *New Media & Society* (2021). [中文题名: 数据与石油: 隐喻、物质性与代谢断裂。]
- . 'AirPods and the Earth: Digital Technologies, Planned Obsolescence and the Capitalocene', *Environment and Planning E: Nature and Space*, 2022. [中文题名: AirPods与地球: 数字技术、计划性报废与资本世。]
- Ursin, Frank, Timmermann, Cristian, Orzechowski, Marcin, and Steger, Florian. 'Diagnosing Diabetic Retinopathy with Artificial Intelligence: What Information Should Be Included to Ensure Ethical Informed Consent?' [In English]. Original Research. *Frontiers in Medicine* 8.1108 (2021). 1–6. [中文题名: 使用人工智能诊断糖尿病视网膜病变: 应纳入哪些信息以确保合乎伦理的知情同意? [英语; 原创研究。]。]
- U.S. Energy Information Administration. *What is U.S. Electricity Generation by Energy Source?*, June 2022, <https://www.eia.gov/tools/faqs/faq.php?id=427&t=3>. [中文题名: 美国按能源来源划分的发电构成是什么?。]
- van Wynsberghe, Aimee. 'Sustainable AI: AI for Sustainability and the Sustainability of AI', *AI and Ethics* (2021): 1–6. [中文题名: 可持续AI: 促进可持续性的AI与AI自身的可持续性。]
- Vaughan, Naomi E., and Gough, Clair. 'Expert Assessment Concludes Negative Emissions Scenarios May Not Deliver', *Environmental Research Letters* 11.9 (2016): 095003. [中文题名: 专家评估得出结论: 负排放情景可能无法兑现。]
- Vonk, Lisa. 'Paying Attention to Waste: Apple's Circular Economy', *Continuum* 32.6 (2018): 745–57. [中文题名: 关注废弃物: Apple的循环经济。]
- Vopson, Melvin M. 'The Information Catastrophe', *AIP Advances* 10.8 (2020): 085014. [中文题名: 信息灾难。]
- Wagner, Ben. 'Ethics as an Escape from Regulation. From 'Ethics-Washing' to Ethics-Shopping?' in *Being Profiled*, Amsterdam: Amsterdam University Press, 2018, pp. 84–9. [中文题名: 作为逃避监管手段的伦理: 从伦理洗白到伦理选购? ; 载《被画像》。]
- Walker, Peter A. 'Political Ecology: Where is the Politics?.' *Progress in Human geography* 31.3 (2007): 363–69. [中文题名: 政治生态学: 政治在哪里?。]

Wiedmann, Thomas O., Schandl, Heinz, Lenzen, Manfred, Moran, Daniel, Suh, Sangwon, West, James, and Kanemoto, Keiichiro. 'The Material Footprint of Nations', proceedings of the National Academy of Sciences 112.20 (2015): 6271–6. [中文题名：各国的物质足迹。]

Zhang, John, and Hoshino, Kazunori. *Molecular Sensors and Nanodevices* 2nd edition. Cambridge: Academic Press, 2019. [中文题名：分子传感器与纳米器件（第2版）。]

原注

- [1] 'The World's Most Valuable Resource Is No Longer Oil, but Data', *The Economist*, 6 May 2017, <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>. [中文题名：世界上最有价值的资源不再是石油，而是数据。] 返回
- [2] Nick Srnicek, *Platform Capitalism*, Cambridge: Polity, 2016. [中文题名：平台资本主义。] 返回
- [3] Jathan Sadowski, 'When Data Is Capital: Datafication, Accumulation, and Extraction', *Big Data & Society* 6.1 (2019): 2053951718820549. [中文题名：当数据成为资本：数据化、积累与攫取。] 返回
- [4] Nick Couldry and Ulises A Mejias, *The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism*, Stanford: Stanford University Press, 2019; Monique Mann and Angela Daly, '(Big) Data and the North-in-South: Australia's Informational Imperialism and Digital Colonialism', *Television & New Media* 20.4 (2019): 379–95. [中文题名：连接的代价：数据如何殖民人类生活并将其据为资本主义所有；（大）数据与南方中的北方：澳大利亚的信息帝国主义和数字殖民主义。] 返回
- [5] Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St. Martin's Press, 2018; Bridle, James, *New Dark Age: Technology and the End of the Future*, London: Verso Books, 2018. [中文题名：自动化不平等：高科技工具如何给穷人画像、管控和惩罚穷人；新黑暗时代：技术与未来的终结。] 返回
- [6] Safiya Umoja Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism*, New York: NYU Press, 2018; Thao Phan and Scott Wark, 'What Personalisation Can Do For You! Or: How to Do Racial Discrimination Without "Race"', *Culture Machine* 20 (2021): 1–29. [中文题名：压迫的算法：搜索引擎如何强化种族主义；个性化能为你做什么！或者：如何在没有“种族”的情况下进行种族歧视。] 返回
- [7] Bernard Stiegler, *Automatic Society: The Future of Work*, Cambridge: Polity, 2017; Mark Andrejevic, *Automated Media*, New York: Routledge, 2020. [中文题名：自动化社会：工作的未来；自动化媒体。] 返回
- [8] Angela Daly, Thilo Hagendorff, Li Hui, Monique Mann, Vidushi Marda, Ben Wagner, Wayne Wei Wang, 'AI, Governance and Ethics: Global Perspectives' in Oreste Pollicino and Giovanni de Gregorio (eds) *Constitutional Challenges in the Algorithmic Society*, Cambridge: Cambridge University Press, 2021. [中文题名：AI、治理与伦理：全球视角；载《算法社会中的宪制挑战》。] 返回

- [9] Andreas Malm and Alf Hornborg, 'The Geology of Mankind? A Critique of the Anthropocene Narrative', *The Anthropocene Review* 1.1 (2014): 62–9; Jason W. Moore, *Anthropocene or Capitalocene?: Nature, History, and the Crisis of Capitalism*, Oakland: Pm Press, 2016; Christophe Bonneuil and Fressoz, Jean-Baptiste, *The Shock of the Anthropocene: The Earth, History and us*, London: Verso Books, 2016. [中文题名: 人类的地质学? 对人类世叙事的批判; 人类世还是资本世? 自然、历史与资本主义危机; 人类世的冲击: 地球、历史与我们。] 返回
- [10] United Nations, 'Paris agreement' in *Report of the Conference of the Parties to the United Nations Framework Convention on Climate Change*, 21st Session, Paris: 2015, https://unfccc.int/files/meetings/paris_nov_2015/application/pdf/paris_agreement_english_.pdf. [中文题名: 巴黎协定; 载《联合国气候变化框架公约缔约方大会第二十一届会议报告》。] 返回
- [11] Eric Nost and Emma Colven, 'Earth for AI: A Political Ecology of Data-Driven Climate Initiatives', *Geoforum* 130 (2022): 23–34. [中文题名: 为AI而存在的地球: 数据驱动气候行动的政治生态学。] 返回
- [12] Alf Hornborg, 'The Political Ecology of the Technocene' in Clive Hamilton, François Gemenne, and Christophe Bonneuil (eds) *The Anthropocene and the Global Environmental Crisis: Rethinking Modernity in a New Epoch*, London, Routledge, 2015, pp. 177–83; Laura Bedford, Monique Mann, Reece Walters, and Marcus Foth, 'A Post-capitalocentric Critique of Digital Technology and Environmental Harm: New Directions at the Intersection of Digital and Green Criminology', *Beyond Cybercrime: New Perspectives on Crime, Harm and Digital Technologies*, *International Journal for Crime, Justice and Social Democracy (Special Issue)*, 11.1 (2022): 167–81. [中文题名: 技术世的政治生态学; 载《人类世与全球环境危机: 在新纪元重新思考现代性》; 对数字技术与环境伤害的后资本中心主义批判: 数字犯罪学与绿色犯罪学交叉领域的新方向; 载《超越网络犯罪: 犯罪、伤害与数字技术的新视角》。] 返回
- [13] Sy Taffel, 'Data and Oil: Metaphor, Materiality and Metabolic Rifts', *New Media & Society* (2021); Eric Nost and Jenny Elaine Goldstein, 'A Political Ecology of Data', *Environment and Planning E: Nature and Space* (2021). [中文题名: 数据与石油: 隐喻、物质性与代谢断裂; 数据的政治生态学。] 返回
- [14] Vasilis Kostakis, Andreas Roos, and Michel Bauwens, 'Towards a Political Ecology of the Digital Economy: Socio-Environmental Implications of Two Competing Value Models', *Environmental Innovation and Societal Transitions* 18 (2016): 82–100. [中文题名: 走向数字经济的政治生态学: 两种竞争价值模式的社会环境含义。] 返回
- [15] Graham Pickren, 'Geographies of E - Waste: Towards a Political Ecology Approach to E - Waste and Digital Technologies', *Geography Compass* 8.2 (2014): 111–24. [中文题名: 电子废弃物地理学: 走向电子废弃物与数字技术的政治生态学进路。] 返回
- [16] Raymond L. Bryant and Lucy Jarosz, 'Ethics in Political Ecology: A Special Issue of Political Geography: Introduction: Thinking About Ethics in Political Ecology', *Political Geography* 23. 7 (2004): 807–12; Peter A. Walker 'Political ecology: where is the politics?.' *Progress in Human Geography* 31.3 (2007): 363–9. [中文题名: 政治生

态学中的伦理：《政治地理学》专刊导论——思考政治生态学中的伦理；政治生态学：政治在哪里？。】 [返回](#)

- [17] Angela Daly, S. Kate Devitt, Monique Mann, ‘AI Ethics Needs Good Data’ In. Pieter Verdegem (ed), *AI for Everyone? Critical Perspectives*, London: University of Westminster Press, 2021. [中文题名：AI伦理需要良好数据；载《人人共享AI？批判性视角》。】 [返回](#)
- [18] Emma Strubell, Ananya Ganesh, and Andrew McCallum, ‘Energy and Policy Considerations for Deep Learning in Nlp’ , arXiv preprint arXiv:1906.02243, 2019; David Patterson, Joseph Gonzalez, Quoc Le, Chen Liang, Lluís-Miquel Munguia, Daniel Rothchild, David So, Maud Texier, and Jeff Dean; ‘Carbon Emissions and Large Neural Network Training’ , arXiv preprint arXiv:2104.10350, 2021. [中文题名：自然语言处理深度学习的能源与政策考量；碳排放与大型神经网络训练。】 [返回](#)
- [19] Roy Schwartz, Jesse Dodge, Noah A. Smith, and Oren Etzioni, ‘Green AI’ , *Communications of the ACM* 63.12 (2020): 54–63; Aimee van Wynsberghe, ‘Sustainable AI: AI for Sustainability and the Sustainability of AI’ , *AI and Ethics* (2021): 1–6. [中文题名：绿色AI；可持续AI：促进可持续性的AI与AI自身的可持续性。】 [返回](#)
- [20] Mél Hogan and Asta Vonderau, ‘The Nature of Data Centers’ , *Culture Machine* 18 (2019): 1–4 Brett Neilson and Ned Rossiter, ‘Automating Labour and the Spatial Politics of Data Centre Technologies’ in Mascha Will-Zocholl and Caroline Roth-Ebner (eds) *Topologies of Digital Work*, London: Palgrave Macmillan Cham, 2021, pp. 77–101. [中文题名：数据中心的自然；劳动自动化与数据中心技术的空间政治；载《数字工作的拓扑》。】 [返回](#)
- [21] Polly Higgins, *Eradicating Ecocide: Exposing the Corporate and Political Practices Destroying the Planet and Proposing the Laws to Eradicate Ecocide*, London: Shepard-Walwyn, 2016. [中文题名：根除生态灭绝：揭露摧毁地球的企业与政治实践，并提出根除生态灭绝的法律。】 [返回](#)
- [22] Ben Wagner, ‘Ethics as an Escape from Regulation. From “Ethics-Washing” to Ethics-Shopping?’ in Emre Bayamlioglu, Irina Baraliuc, Liisa Albertha Wilhelmina Janssens, and Mireille Hildebrandt (eds) *Being Profiled*, Amsterdam: Amsterdam University Press, 2018, 84–9; Yochai Benkler, ‘Don’t Let Industry Write the Rules for AI’ , *Nature*, 569.7754 (2019): 161–2. [中文题名：作为逃避监管手段的伦理：从伦理洗白到伦理选购？；载《被画像》；别让行业制定AI规则。】 [返回](#)
- [23] Alexandre Lacoste, Alexandra Luccioni, Victor Schmidt, and Thomas Dandres, ‘Quantifying the Carbon Emissions of Machine Learning’ (preprint) arXiv: 1910.09700, 2019; Roy Schwartz, Jesse Dodge, Noah A. Smith, and Oren Etzioni, ‘Green AI’ ; E.M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell, ‘On the Dangers of Stochastic Parrots: Can Language Models Be Too Big??’ *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 1 March 2021: 610–23. [中文题名：量化机器学习的碳排放；绿色AI；论随机鹦鹉的危险：语言模型是否可能过于庞大？。】 [返回](#)

- [24] Emma Strubell, Ananya Ganesh, and Andrew McCallum, ‘Energy and Policy Considerations for Deep Learning in Nlp’ . [中文题名：自然语言处理深度学习的能源与政策考量。] 返回
- [25] Apple, *iPhone 13 Product Environmental Report*, 14 September 2021, https://www.apple.com/environment/pdf/products/iphone/iPhone_13_PER_Sept2021.pdf. [中文题名：iPhone 13产品环境报告。] 返回
- [26] 使用专为机器学习设计的张量处理器（TPU）取代图形处理器（GPU），使训练时间从274,120小时降至32,623小时。不过，由于缺乏TPU耗能的公开资料，文中只提供了用GPU进行训练的CO2e数值。返回
- [27] David Patterson, Joseph Gonzalez, Quoc Le, Chen Liang, Lluís-Miquel Munguia, Daniel Rothchild, David So, Maud Texier, and Jeff Dean, ‘Carbon Emissions and Large Neural Network Training’ . [中文题名：碳排放与大型神经网络训练。] 返回
- [28] Emma Strubell, Ananya Ganesh, and Andrew McCallum, ‘Energy and Policy Considerations for Deep Learning in Nlp’ . [中文题名：自然语言处理深度学习的能源与政策考量。] 返回
- [29] George Leopold, ‘AWS to Offer Nvidia’s T4 Gpus for AI Inferencing’ , *HPC Wire*, 19 March 2019, <https://www.hpcwire.com/2019/03/19/aws-upgrades-its-gpu-backed-ai-inference-platform/>. [中文题名：AWS将提供Nvidia的T4 GPU用于AI推理。] 返回
- [30] Jeff Barr, *Amazon EC2 Update—Inf1 Instances with AWS Inferentia Chips for High Performance Cost-Effective Inferencing*, 3 December 2019, <https://aws.amazon.com/blogs/aws/amazon-ec2-update-inf1-instances-with-aws-inferentia-chips-for-high-performance-cost-effective-inferencing/>. [中文题名：Amazon EC2更新——配备AWS Inferentia芯片、实现高性能与高成本效益推理的Inf1实例。] 返回
- [31] Eric Masanet, Arman Shehabi, Nuoa Lei, Sarah Smith, and Jonathan Koomey, ‘Recalibrating Global Data Center Energy-Use Estimates’ , *Science* 367.6481 (2020): 984–6. [中文题名：重新校准全球数据中心能耗估算。] 返回
- [32] S.E., Andrae Anders, ‘Hypotheses for Primary Energy Use, Electricity Use and CO2 Emissions of Global Computing and its Shares of the Total Between 2020 and 2030’ , *WSEAS Transactions on Power Systems*, 15 (2020): 50–9; Rabih Bashroush and Andy Lawrence, *Beyond PUE: Tackling Its Wasted Terawatts*, Uptime Institute, January 2020, <https://uptimeinstitute.com/beyond-pue-tackling-it%E2%80%99s-wasted-terawatts>. [中文题名：关于2020至2030年全球计算的一次能源使用、电力使用和CO2排放及其占总量比例的假设；超越PUE：应对其浪费的太瓦级电力。] 返回
- [33] Anders, ‘Hypotheses for Primary Energy Use, Electricity Use and CO2 Emissions of Global Computing and its Shares of the Total Between 2020 and 2030’ . [中文题名：关于2020至2030年全球计算的一次能源使用、电力使用和CO2排放及其占总量比例的假设。] 返回
- [34] Synergy Research Group, ‘Hyperscale Data Center Count Grows to 659—ByteDance Joins the Leading Group’ , 13 September 2021, <https://www.srgresearch.com/articles/hyperscale-data-center-count-grows-to-659-bytedance-joins-the-leading-group>. [中文题名：超大规模数据中心数量增至659个——字节跳动加入领先群体。] 返回

- [35] Dan Swinhoe, 'Microsoft, Amazon, and Google Operate Half the World' s 600 Hyperscale Data Centers' , 2021, <https://www.datacenterdynamics.com/en/news/microsoft-amazon-and-google-operate-half-the-worlds-600-hyperscale-data-centers/>. [中文题名: Microsoft、Amazon与Google运营着全球600个超大规模数据中心中的一半。] 返回
- [36] Swinhoe, 'Microsoft, Amazon, and Google Operate Half the World' s 600 Hyperscale Data Centers' . [中文题名: Microsoft、Amazon与Google运营着全球600个超大规模数据中心中的一半。] 返回
- [37] Nick Srnicek, *Platform Capitalism*, Cambridge: Polity, 2016. [中文题名: 平台资本主义。] 返回
- [38] Arman Shehabi, Sarah J. Smith, Eric Masanet, and Jonathan Koomey, 'Data Center Growth in the United States: Decoupling the Demand for Services from Electricity Use' , *Environmental Research Letters* 13.12 (2018): 124030. [中文题名: 美国数据中心增长: 服务需求与用电脱钩。] 返回
- [39] Mél Hogan, 'Data Flows and Water Woes: The Utah Data Center' , *Big Data & Society* 2.2: (2015): 1-12. [中文题名: 数据流与水之忧: 犹他数据中心。] 返回
- [40] Nicole Starosielski, 'Thermocultures of Geological Media' , *Cultural Politics* 12.3 (2016): 293-309. [中文题名: 地质媒介的热文化。] 返回
- [41] David Mytton, 'Data Centre Water Consumption' , *Npj Clean Water* 4.1 (2021): 11. [中文题名: 数据中心用水量。] 返回
- [42] Arman Shehabi et al., *United States Data Center Energy Usage Report*. U.S. Department of Energy Office of Scientific and Technical Information, 2016, <https://www.osti.gov/biblio/1372902/>. [中文题名: 美国数据中心能源使用报告。] 返回
- [43] Cheryl A. Dieter, Molly A. Maupin, Rodney R. Caldwell, Melissa A. Harris, Tamara I. Ivahnenko, John K. Lovelace, Nancy L. Barber, and Kristin S. Linsey, *Estimated Use of Water in the United States in 2015*, Reston, VA: U.S. Geological Survey, 2018. [中文题名: 2015年美国用水量估算。] 返回
- [44] Md Abu Bakar Siddik, Arman Shehabi, and Landon Marston, 'The Environmental Footprint of Data Centers in the United States' , *Environmental Research Letters* 16.6 (2021): 064017. [中文题名: 美国数据中心的环境足迹。] 返回
- [45] Mél Hogan, 'Big Data Ecologies' , *Ephemera* 18.3 (2018): 631-57. [中文题名: 大数据生态。] 返回
- [46] Roy Schwartz, Jesse Dodge, Noah A. Smith, and Oren Etzioni, 'Green AI' ; Bender, Gebru, McMillan-Major and Shmitchell, 'On the Dangers of Stochastic Parrots: Can Language Models Be Too Big??' . [中文题名: 绿色AI; 论随机鹦鹉的危险: 语言模型是否可能过于庞大? 。] 返回
- [47] Sy Taffel, *Digital Media Ecologies*, New York and London: Bloomsbury, 2019, 162-9. [中文题名: 数字媒介生态。] 返回

- [48] Nicholas Niarchos, ‘The Dark Side of Congo’ s Cobalt Rush’ , *The New Yorker*, 2021, <https://www.newyorker.com/magazine/2021/05/31/the-dark-side-of-congos-cobalt-rush>. [中文题名: 刚果钴矿热潮的阴暗面。] 返回
- [49] Julie Michelle Klinger, *Rare Earth Frontiers*, Ithaca, London: Cornell University Press, 2017. [中文题名: 稀土边疆。] 返回
- [50] Rembrandt H. E. M. Koppelaar and Hendrik Koppelaar, ‘The Ore Grade and Depth Influence on Copper Energy Inputs’ , *BioPhysical Economics and Resource Quality* 1.2 (2016): 11. [中文题名: 矿石品位与埋深对铜生产能源投入的影响。] 返回
- [51] 应当指出, 钴主要是那些以开采铜和镍为主的矿山所产出的副产品。返回
- [52] Beiser Vince, *The World in a Grain: The Story of Sand and How It Transformed Civilization*, New York: Riverhead Books, 2018. [中文题名: 一粒沙中的世界: 沙子的故事及其如何改变文明。] 返回
- [53] John Zhang and Kazunori Hoshino, *Molecular Sensors and Nanodevices* 2nd edition, Cambridge: Academic Press, 2019. [中文题名: 分子传感器与纳米器件 (第2版)。] 返回
- [54] Alf Hornborg, *The Power of the Machine: Global inequalities of Economy, Technology, and Environment*, Lanham: Altamira Press, 2001. [中文题名: 机器的力量: 经济、技术与环境的全球不平等。] 返回
- [55] Nick Couldry and Ulises A Mejias, *The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism*, Stanford: Stanford University Press, 2019. [中文题名: 连接的代价: 数据如何殖民人类生活并将其据为资本主义所有。] 返回
- [56] Alf Hornborg, ‘Towards an Ecological Theory of Unequal Exchange: Articulating World System Theory and Ecological Economics’ , *Ecological Economics* 25.1 (1998): 127–36; Alf Hornborg and Joan Martinez-Alier, ‘Ecologically Unequal Exchange and Ecological Debt’ , *Journal of Political Ecology* 23.1(2016): 328–33. [中文题名: 走向不平等交换的生态理论: 联结世界体系理论与生态经济学; 生态不平等交换与生态债务。] 返回
- [57] Raj Patel and Jason W. Moore, *A History of the World in Seven Cheap Things: A Guide to Capitalism, Nature, and the Future of the Planet*, New York: Verso, 2017; Andreas Malm and Alf Hornborg, ‘The Geology of Mankind? A Critique of the Anthropocene Narrative’ , *The Anthropocene Review* 1.1 (2014): 62–9. [中文题名: 以七种廉价之物书写的世界史: 资本主义、自然与地球未来指南; 人类的地质学? 对人类世叙事的批判。] 返回
- [58] Vanessa Forti, Cornelis P. Balde, Ruediger Kuehr, and Garam Bel, *The Global E-Waste Monitor 2020: Quantities, Flows and the Circular Economy Potential*, Bonn, Geneva and Rotterdam: United Nations University/United Nations Institute for Training and Research, International Telecommunication Union, and International Solid Waste Association, 2020. [中文题名: 2020年全球电子废弃物监测: 数量、流动与循环经济潜力。] 返回

- [59] Christian, Dorninger, Alf Hornborg, David J. Abson, Henrik Von Wehrden, Anke Schaffartzik, Stefan Giljum, John-Oliver Engler, et al., 'Global Patterns of Ecologically Unequal Exchange: Implications for Sustainability in the 21st Century', *Ecological Economics* 179 (2021): 106824. [中文题名: 生态不平等交换的全球模式: 对21世纪可持续性的启示。] 返回
- [60] Katarina Sluis and Fei-Fei Li, *Where Did ImageNet Come From?*, November 2019, <https://unthinking.photography/articles/where-did-imagenet-come-from>. [中文题名: ImageNet从何而来?。] 返回
- [61] Birgitta Bergvall-Kåreborn and Debra Howcroft, 'Amazon Mechanical Turk and the commodification of labour', *New Technology, Work and Employment*, 29.3 (2014): 213–23; Jathan Sadowski, 'The Internet of Landlords: Digital Platforms and New Mechanisms of Rentier Capitalism', *Antipode*, 52 (2020): 562–80. [中文题名: Amazon Mechanical Turk与劳动商品化; 房东的互联网: 数字平台与食利资本主义的新机制。] 返回
- [62] Strubell, Ganesh, and McCallum, 'Energy and Policy Considerations for Deep Learning in NLP'; Schwartz, Dodge, Smith, and Etzioni, 'Green AI'; Patterson, Gonzalez, Le, Liang, Munguia, Rothchild, So, Texier, and Dean, 'Carbon Emissions and Large Neural Network Training'; Aimee van Wynsberghe, 'Sustainable AI: AI for Sustainability and the Sustainability of AI', *AI and Ethics* (2021): 1–6; Bender, Gebru, McMillan-Major, and Shmitchell 'On the Dangers of Stochastic Parrots: Can Language Models Be Too Big??'. [中文题名: 自然语言处理深度学习的能源与政策考量; 绿色AI; 碳排放与大型神经网络训练; 可持续AI: 促进可持续性的AI与AI自身的可持续性; 论随机鹦鹉的危险: 语言模型是否可能过于庞大?。] 返回
- [63] Schwartz, JDodge, Smith, and Etzioni, 'Green AI; van Wynsberghe, 'Sustainable AI: AI for Sustainability and the Sustainability of AI'. [中文题名: 绿色AI; 可持续AI: 促进可持续性的AI与AI自身的可持续性。] 返回
- [64] Thomas O. Wiedmann, Heinz Schandl, Manfred Lenzen, Daniel Moran, Sangwon Suh, James West, and Keiichiro Kanemoto, 'The Material Footprint of Nations', *proceedings of the National Academy of Sciences* 112.20 (2015): 6271–6276; Jason Hickel and Giorgos Kallis, 'Is Green Growth Possible?', *New Political Economy* (2019): 1–18. [中文题名: 各国的物质足迹; 绿色增长可能吗?。] 返回
- [65] 例如参见: Esther Sanyé-Mengual, Michaela Secchi, Sara Corrado, Antoine Beylot, and Serenella Sala, 'Assessing the Decoupling of Economic Growth from Environmental Impacts in the European Union: A Consumption-based Approach', *Journal of Cleaner Production*, 236 (2019): 117535. [中文题名: 评估欧盟经济增长与环境影响的脱钩: 基于消费的进路。] 返回
- [66] Jason Hickel and Giorgos Kallis, 'Is Green Growth Possible?' *New Political Economy* (2019): 1–18. [中文题名: 绿色增长可能吗?。] 返回
- [67] Jason Hickel, *Less Is More: How Degrowth Will Save the World*, New York: Penguin Random House, 2020; Wiedmann, Schandl, Lenzen, Moran, Suh, West, and Kanemoto 'The Material Footprint of Nations'. [中文题名: 少即是多: 去增长将如何拯救世界; 各国的物质足迹。] 返回

- [68] Stefan Bringezu, 'Possible Target Corridor for Sustainable Use of Global Material Resources', *Resources* 4.1 (2015): 25–54. These global figures are very coarse-grained, as differing materials have different impacts, as do the technologies and processes used for material extraction. Still, despite this homogenisation these figures are useful indicators of the rapid growth in material footprint since 1950. [中文题名：全球物质资源可持续使用的可能目标区间。] [原注说明译文：这些全球数据非常粗略，因为不同材料具有不同影响，开采材料所用的技术和工艺也各有不同。不过，尽管进行了这种同质化处理，这些数字仍是有用的指标，能够反映1950年以来物质足迹的快速增长。] 返回
- [69] Hickel *Less Is More* 110. [中文题名：少即是多：去增长将如何拯救世界。] 返回
- [70] Koppelaar and Koppelaar, 'The Ore Grade and Depth Influence on Copper Energy Inputs', 11; Taffel, 'Data and Oil'. [中文题名：矿石品位与埋深对铜生产能源投入的影响；数据与石油：隐喻、物质性与代谢断裂。] 返回
- [71] Anders, 'Hypotheses for primary energy use, electricity use and CO2 emissions of global computing and its shares of the total between 2020 and 2030'. [中文题名：关于2020至2030年全球计算的一次能源使用、电力使用和CO2排放及其占总量比例的假设。] 返回
- [72] Patterson, Gonzalez, Le, Liang, Munguia, Rothchild, So, Texier, and Dean, 'Carbon Emissions and Large Neural Network Training'. [中文题名：碳排放与大型神经网络训练。] 返回
- [73] Gemma A. Brady, Nikil Kapur, Jonathan L. Summers, and Harvey M. Thompson, 'A Case Study and Critical Assessment in Calculating Power Usage Effectiveness for a Data Center', *Energy Conversion and Management* 76 (2013): 155–61. [中文题名：数据中心电能使用效率计算的案例研究与批判性评估。] 返回
- [74] Anders, 'Hypotheses for Primary Energy Use, Electricity Use and CO2 Emissions of Global Computing and its Shares of the Total Between 2020 and 2030', 53. [中文题名：关于2020至2030年全球计算的一次能源使用、电力使用和CO2排放及其占总量比例的假设。] 返回
- [75] Stephan G. Bunker, 'Raw Materials and the Global Economy', *Society and Natural Resources* 9.4 (1996): 419–29. [中文题名：原材料与全球经济。] 返回
- [76] William Stanley Jevons (1865) *The Coal Question*, 3rd edition, New York: Augustus M. Kelley, 1965. [中文题名：煤炭问题。] 返回
- [77] John Bellamy Foster, Brett Clark, and Richard York, 'Capitalism and the Curse of Energy Efficiency', *Monthly Review* 62.6 (2010): 183. [中文题名：资本主义与能源效率的诅咒。] 返回
- [78] Evgeny Morozov, *To Save Everything, Click Here: The Folly of Technological Solutionism*, London: Penguin, 2014. Our use of technological solutionism is related to the term's popularisation by Morozov, insofar as it demarcates a logic unique to 21st-century digital capitalism which seeks to resolve problems through optimization, efficiency and market-based computational technologies. Morozov productively outlines how solutionism exemplifies a desire to solve problems that do not

meaningfully exist, as epitomized by smart fridges or bin cameras. While we concur with the critique of these technologies as needlessly extending corporate dataveillance, our critique of technological solutionism instead addresses how this logic adversely impacts contemporary responses to urgent ecological and social justice issues. [中文题名: 要拯救一切, 请点击这里: 技术解决主义的愚妄。] [原注说明译文: 我们对“技术解决主义”的使用, 与Morozov使这一术语普及的方式相关, 因为它标识了一种21世纪数字资本主义特有的逻辑, 试图通过优化、效率和基于市场的计算技术来解决问题。Morozov富有启发性地指出, 解决主义体现了一种解决实际上并无实质意义的问题的欲望, 智能冰箱或垃圾桶摄像头就是典型例子。我们赞同对这些技术的批评, 即它们不必要地扩展了企业的数据监控; 但我们的技术解决主义批判, 转而讨论这种逻辑如何对当代针对紧迫生态与社会正义问题的回应产生不利影响。] 返回

- [79] ‘PM Defends his “Technology Not Taxes” Approach to Cutting Emissions’, *ABC News*, 23 April 2021, <https://www.abc.net.au/news/2021-04-23/pm-defends-his-technology-not-taxes-approach-to-emissions/13315616>. [中文题名: 总理为其“技术而非税收”的减排路径辩护。] 返回

- [80] George W. Bush and John Howard, *Joint Statement on Climate Change and Energy by President Bush and Australian Prime Minister John Howard*, 4 September 2007, <https://georgewbush-whitehouse.archives.gov/news/releases/2007/09/20070904-8.html>. [中文题名: 布什总统与澳大利亚总理John Howard关于气候变化和能源的联合声明。] 返回

- [81] Kevin Anderson and Glen Peters, ‘The Trouble with Negative Emissions’, *Science* 354. 6309 (2016): 182–3. [中文题名: 负排放的问题。] 返回

- [82] Declan Kuch, ‘“Fixing” Climate Change through Carbon Capture and Storage: Situating Industrial Risk Cultures’, *Futures* 92 (2017): 90–9. [中文题名: 通过碳捕集与封存“修复”气候变化: 定位工业风险文化。] 返回

- [83] IPCC, *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Core Writing Team, Rajendra Pachauri and Leo Meyer (eds)]. IPCC, Geneva: Switzerland, 2014. [中文题名: 气候变化2014: 综合报告——第一、第二和第三工作组对政府间气候变化专门委员会第五次评估报告的贡献。] 返回

- [84] Hickel, *Less Is More*. [中文题名: 少即是多: 去增长将如何拯救世界。] 返回

- [85] Anderson and Peters, ‘The Trouble with Negative Emissions’, 182–3; Naomi E. Vaughan and Clair Gough, ‘Expert Assessment Concludes Negative Emissions Scenarios May Not Deliver’, *Environmental Research Letters* 11.9 (2016). [中文题名: 负排放的问题; 专家评估得出结论: 负排放情景可能无法兑现。] 返回

- [86] Anderson and Peters, Glen, ‘The Trouble with Negative Emissions’, 183. [中文题名: 负排放的问题。] 返回

- [87] Planting trees has also been criticized for employing a reductionist framing that can reduce biodiversity, increase deforestation, deflect attention from phasing out fossil fuels (see Karen D. Holl, and Pedro H. S. Brancalion, ‘Tree Planting Is Not a Simple Solution’, *Science* 368.6491 (2020): 580–1; Nathalie Seddon, et al. ‘Getting the

- Message Right on Nature-Based Solutions to Climate Change, *Global Change Biology* 27. 8 (2021): 1518–46.) and may not result in net carbon sequestration on decadal timescales (see Nina L. Friggens, Alison J. Hester, Ruth J. Mitchell, Thomas C. Parker, Jens-Arne Subke, and Philip A. Wookey, ‘Tree Planting in Organic Soils Does Not Result in Net Carbon Sequestration on Decadal Timescales’, *Global Change Biology* 26.9 (2020): 5178–88). [中文题名: 植树并非简单解决方案; 准确传达以自然为基础的气候变化解决方案的信息; 在有机土壤中植树不会在十年尺度上产生净碳封存。] [原注说明译文: 植树也受到批评, 因为它采用还原论框架, 可能降低生物多样性、加剧森林砍伐, 并转移人们对逐步淘汰化石燃料的关注 (见Holl与Brancaion; Seddon等)。此外, 植树在十年尺度上可能并不产生净碳封存 (见Friggens等)。] 返回
- [88] Sy Taffel, ‘Hopeful Extinctions? Tesla, Technological Solutionism and the Anthropocene’, *Culture Unbound: Journal of Current Cultural Research* 10.2 (2018): 163–84. [中文题名: 充满希望的灭绝? Tesla、技术解决主义与人类世。] 返回
- [89] Rianne Riems, ‘Decoupling as Rhetorical Strategy in Google’s Green Discourse’, *AoIR Selected Papers of Internet Research*, 2021. [中文题名: 作为Google绿色话语修辞策略的脱钩。] 返回
- [90] Google, ‘Google Environmental Report 2020’, 2021, <https://www.gstatic.com/gumdrop/sustainability/google-2019-environmental-report.pdf> [中文题名: Google 2020年环境报告。] 返回
- [91] Janet Ranganathan, Laurent Corbier, P. Bhatia, Simon Schmitz, Peter Gage, and Kjell Oren, ‘The Greenhouse Gas Protocol: A Corporate Accounting and Reporting Standard (Revised Edition)’, Washington, DC: World Resources Institute and World Business Council for Sustainable Development, 2004. [中文题名: 温室气体核算体系: 企业核算与报告标准 (修订版)。] 返回
- [92] Google的环境数据没有把范围3排放进一步分类, 但Apple的2021年《环境进展报告》(2021b, 第67页)作了分类, 发现该公司总计22,600,000吨CO₂e的碳足迹中, 制造环节占16,100,000吨, 即总量的71%。 返回
- [93] Google, ‘Google Environmental Report 2020’, 2021, <https://www.gstatic.com/gumdrop/sustainability/google-2019-environmental-report.pdf>. [中文题名: Google 2020年环境报告。] 返回
- [94] Netto de Freitas, Sebastião Vieira, Marcos Felipe Falcão Sobral, Ana Regina Bezerra Ribeiro, and Gleibson Robert da Luz Soares, ‘Concepts and Forms of Greenwashing: A Systematic Review’, *Environmental Sciences Europe*, 32.19 (2020). [中文题名: 漂绿的概念与形式: 系统综述。] 返回
- [95] Apple, *Environmental Progress Report*, April 2021, https://www.apple.com/environment/pdf/Apple_Environmental_Progress_Report_2021.pdf. [中文题名: 环境进展报告。] 返回
- [96] U.S. Energy Information Administration, *Electricity Explained, Electricity in the United States*, July 2022, <https://www.eia.gov/energyexplained/electricity/electricity-in-the-us.php>. [中文题名: 电力释义: 美国的电力。] 返回

- [97] Apple has also been active in challenging right-to-repair movements and engaging in strategies of planned obsolescence to drive ongoing consumption of its devices leading to more mineral extraction and waste creation. See Lisa Vonk, 'Paying Attention to Waste: Apple's Circular Economy', *Continuum* 32.6 (2018): 745–57; Bedford, Mann, Walters, and Foth, 'A post-Capitalocentric Critique of Digital Technology and Environmental Harm: New Directions at the Intersection of Digital and Green Criminology'; Sy Taffel, 'AirPods and the Earth: Digital Technologies, Planned Obsolescence and the Capitalocene', *Environment and Planning E: Nature and Space*, 2022. [中文题名: 关注废弃物: Apple的循环经济; 对数字技术与环境伤害的后资本中心主义批判: 数字犯罪学与绿色犯罪学交叉领域的新方向; 载《超越网络犯罪: 犯罪、伤害与数字技术的新视角》; AirPods与地球: 数字技术、计划性报废与资本世。] [原注说明译文: Apple还积极反对维修权运动, 并采取计划性报废策略, 推动消费者持续购买其设备, 从而导致更多矿物开采与废弃物产生。参见所列Vonk、Bedford等与Taffel的研究。] 返回
- [98] Hannah Ritchie and Max Roser, *Energy Mix*, 2021, <https://ourworldindata.org/energy-mix>. [中文题名: 能源组合。] 返回
- [99] Laura Bedford, Monique Mann, Reece Walters, and Marcus Foth, 'A Post-capitalocentric Critique of Digital Technology and Environmental Harm: New Directions at the Intersection of Digital and Green Criminology in 'Beyond Cybercrime: New perspectives on crime, harm and digital technologies'', *International Journal for Crime, Justice and Social Democracy (Special Issue)*, 11.1 (2022): 167–81. [中文题名: 对数字技术与环境伤害的后资本中心主义批判: 数字犯罪学与绿色犯罪学交叉领域的新方向; 载《超越网络犯罪: 犯罪、伤害与数字技术的新视角》。] 返回
- [100] Aaron Bastani, *Fully Automated Luxury Communism*, London: Verso Books, 2020. [中文题名: 全自动奢华共产主义。] 返回
- [101] Thilo Hagendorff, 'The Ethics of AI Ethics: An Evaluation of Guidelines', *Minds & Machines* 30, 2020: 99–120. [中文题名: AI伦理的伦理: 指南评估。] 返回
- [102] Manu V. Mathai, Cindy Isenhour, Dimitris Stevis, Philip Vergragt, Magnus Bengtsson, Sylvia Lorek, Lars Fogh Mortensen et al., 'The Political Economy of (un) Sustainable Production and Consumption: A Multidisciplinary Synthesis for Research and Action', *Resources, Conservation and Recycling* 167 (2021): 105265. [中文题名: (不) 可持续生产与消费的政治经济学: 面向研究与行动的多学科综合。] 返回
- [103] David Harvey, *The Enigma of Capital and the Crises of Capitalism*, London: Profile Books, 2010. [中文题名: 资本之谜与资本主义危机。] 返回
- [104] Giorgos Kallis, *Degrowth*, Newcastle-Upon-Tyne: Agenda Publishing, 2018; Brian Massumi, *99 Theses on the Revaluation of Value: A Postcapitalist Manifesto*, Minneapolis: University of Minnesota Press, 2018; Alf Hornborg, 'Energy, Space, and Movement: Toward a Framework for Theorizing Energy Justice', *Geografiska Annaler: Series B, Human Geography* 102.1 (2020): 8–20; Hickel, *Less Is More*; Tim Jackson, *Post Growth: Life After Capitalism*, Cambridge: Polity, 2020. [中文题名: 去增

- 长；关于价值重估的99条论纲：后资本主义宣言；能源、空间与运动：迈向能源正义理论化的框架；少即是多：去增长将如何拯救世界；后增长：资本主义之后的生活。】 返回
- [105] Thomas Piketty, *Capital in the Twenty-first Century*, Cambridge: Harvard University Press, 2018. [中文题名：21世纪资本论。】 返回
- [106] 例如参见，Paul Robbins, ‘Is Less More...or Is More Less? Scaling the Political Ecologies of the Future’ , *Political Geography* 76 (2020): 102018; Erik Gómez-Baggethun, ‘More Is More: Scaling Political Ecology within Limits to Growth’ , *Political Geography* 76 (2020): 102095. [中文题名：少即是多……还是多即是少？未来政治生态学的尺度拓展；更多就是更多：在增长极限内拓展政治生态学的尺度。】 返回
- [107] 例如参见，Joel Kovel, *The Enemy of Nature: The End of Capitalism or the End of the World?*, London: Zed Books, 2007. [中文题名：自然之敌：资本主义的终结还是世界的终结？。】 返回
- [108] Hickel, *Less Is More*. [中文题名：少即是多：去增长将如何拯救世界。】 返回
- [109] Jathan Sadowski, Salomé Viljoen, and Meredith Whittaker, ‘Everyone Should Decide How Their Digital Data are Used—Not Just Tech Companies’ , *Nature* 595 (2021): 169–71. [中文题名：人人都应决定自己的数字数据如何使用，而不仅仅是科技公司。】 返回
- [110] Srnicek, *Platform Capitalism*; James Muldoon, *Platform Socialism: How to Reclaim our Digital Future from Big Tech*, London: Pluto Press, 2022. [中文题名：平台资本主义；平台社会主义：如何从大型科技企业手中夺回我们的数字未来。】 返回
- [111] Christian Kerschner, Petra Wächter, Linda Nierling, and Melf-Hinrich Ehlers, ‘Degrowth and Technology: Towards Feasible, Viable, Appropriate and Convivial Imaginaries’ , *Journal of Cleaner Production* 197 (2018): 1619–36; Beinsteiner, Andreas, ‘Conviviality, the Internet, and AI. Ivan Illich, Bernard Stiegler, and the Question Concerning Information-technological Self-limitation’ , *Open Cultural Studies* 4.1 (2020): 131–42. [中文题名：去增长与技术：迈向可行、可存续、适切且共生的想象；共生性、互联网与AI：Ivan Illich、Bernard Stiegler以及信息技术自我限制问题。】 返回
- [112] Vasilis Kostakis, Kostas Latoufis, Minas Liarokapis, and Michel Bauwens, ‘The Convergence of Digital Commons with Local Manufacturing From a Degrowth Perspective: Two Illustrative Cases’ , *Journal of Cleaner Production* 197 (2018): 1684–93. [中文题名：从去增长视角看数字共有资源与本地制造的融合：两个说明性案例。】 返回
- [113] Albert-László Barabási and Albert, Réka, ‘Emergence of Scaling in Random Networks’ , *Science*, 286.5439 (1999): 509–12; Srnicek, *Platform Capitalism*. [中文题名：随机网络中标度的涌现；平台资本主义。】 返回
- [114] Nicholas Garnham, ‘Editorial’ , *Media, Culture & Society* 14.3 (1992): 339–42. [中文题名：社论。】 返回
- [115] Jathan Sadowski, ‘When Data Is Capital: Datafication, Accumulation, and Extraction’ , *Big Data & Society* 6.1 (2019): 1–12. [中文题名：当数据成为资本：数据化、积累与攫取。】 返回

- [116] Nick Couldry and Ulises A. Mejias, 'Data Colonialism: Rethinking Big Data' s Relation to the Contemporary Subject' , *Television & New Media* 20.4 (2018): 336–49; Monique Mann and Angela Daly, '(Big) Data and the North-in-South: Australia' s Informational Imperialism and Digital Colonialism' , *Television & New Media* 20.4 (2019): 379–95. [中文题名：数据殖民主义：重新思考大数据与当代主体的关系；（大）数据与南方中的北方：澳大利亚的信息帝国主义和数字殖民主义。] [返回](#)
- [117] Tahu Kukutai and John Taylor (eds), *Indigenous Data Sovereignty: Toward an Agenda*. Canberra: ANU Press, 2016. [中文题名：原住民数据主权：迈向一项议程。] [返回](#)
- [118] Monique Mann, Peta Mitchell, Marcus Foth, and Irina Anastasiu, '#BlockSidewalk to Barcelona: Technological Sovereignty and the Social Licence to Operate Smart Cities' , *Journal of the Association for Information Science and Technology* 71.9 (2020): 1103–15. [中文题名：从#BlockSidewalk到巴塞罗那：技术主权与智慧城市运营的社会许可。] [返回](#)

日常AI伦理： 通过人脸识别从全球走向地方

Angela Daly

引言

有关AI伦理框架及其他倡议的重要讨论，多发生在国际或国家层面；尤其是从人权进路出发的讨论，可能声称具有普遍或全球性的适用范围与意义。^[1]在北美、欧洲和东亚等地区的主要国家之外，^[2]国家层面的AI伦理讨论与活动较少受到关注；即使在这些“主要国家”内部，次国家层面——例如获授权的地区或省级行政机构——及地方层面的讨论与活动，也往往被忽略，让位于据称影响更大的“更高层级”讨论。然而，这构成了一个问题，因为除非理解人们如何在地方层面遭遇AI、与之协商并对其提出争议，否则这些高层级讨论并没有多少意义。^[3]

即使在这些主要国家和地区内部，更地方性的AI伦理讨论与实践，也可能被研究者忽视，或被认为相关性影响力较低，对政策制定者和企业而言，甚至可能令人不便。以我目前所在、也是本章讨论对象的英国为例，学术界的“影响”是指“优秀研究对社会和经济作出的、可被证明的贡献”。^[4]英国的研究影响政策，导致批评政府政策的研究比其他政策相关研究得分更低；一些学者认为，它“使研究资助偏向政治意识形态与大企业的利益”。^[5]某些AI伦理活动在国家和国际层面看似重要，这似乎也吸引了研究经费及其他形式的资助，至少部分是以这种被假定的“影响力”为根据，并形成随之而来的“美德经济”：AI伦理由大型科技企业资助，产出供大型科技企业消费的成果。^[6]

确实，虽然我主持着英国研究与创新署（UKRI）资助的自动化及AI主题研究项目，但本文是在“无资助”基础上写成的，因为它不属于那些项目的范围。UKRI是英国的公共研究资助机构，却在政策指引下高度强调“商业化”；正如 Finn 所言，这些政策导致“英国本土创新的商品化”。^[7]我关于人脸识别与苏格兰的其他工作，也是在大学雇主分配给我的非外部资助研究时间，以及正式工作时间之外的个人时间完成的。我认为，这说明了学术研究竞争性资助的优先次序：在苏格兰这样更地方性的情境中批判性研究人脸识别，不如旨在促进AI和自动化用于英国医疗与制造业的研究有吸引力——后者正是我获得资助的研究。这一认识补充了本书其他作者提出的因素，例如企业优先事项、政府优先事项，以及资助获得者之间的性别不平衡，而且很可能还有其他不平衡。^[8]不过，本文也印证了 Edwards 的观点：无资助研究是“一个可以面对并处理由朝不同方向拉扯的学术身份形式所产生张力的空间”。^[9]就我而言，这意味着我获得了机会，能够“对知识作出创造性且

由求知驱动的贡献”，^[10]并抵抗我自己在争取AI研究经费方面的新自由主义式成功！

本文也批判性地审视英国的AI伦理。如上所述，批评英国政府政策的研究，在研究影响评分上可能低于其他政策导向研究。英国政府大力投资AI，包括治理和政策方面，直接或间接支持了一群行动者及项目，例如 Alan Turing Institute、Digital Catapult 和 Centre for Data Ethics and Innovation。Ada Lovelace Institute 表面上虽然独立，却是与包括 Alan Turing Institute 在内的若干英国政府资助机构“合作”成立的，也接受了UKRI的资助。英国作为民族国家，积极参与全球AI治理讨论；在国内也制定了自身的《国家AI战略》，最近又发布政策文件，阐述其“促进创新的AI监管进路”，并在这一过程中避开具有法律约束力的规范。^[11]英国当前AI进路由若干主题支撑，包括优先考虑“创新”，以及与欧盟数据保护进路分道扬镳、转而更接近美国；两者都与英国脱欧后的地缘政治及经济立场有关。^[12]Ossewaarde 和 Gulenc 发现，英国的AI进路具有数字乌托邦和技术解决主义色彩，借助英国帝国主义历史及其在工业革命中的领导地位，把脱欧后的英国投射为未来AI领域的新帝国式“世界领导者”，同时粉饰了AI可能削弱民主的“阴暗面”。^[13]

此外，Ossewaarde 和 Gulenc 指出，英国AI政策带有强烈的技术官僚特征。^[14]参与AI治理和伦理讨论、制定各种原则或规则的那些人，往往是“技术导向”的专家，与普通入及其经验相去甚远，使AI治理迄今成为一项“精英主义工程”。^[15]AI伦理也“主要由男性塑造”，呈现更普遍的“缺乏多样性”，^[16]而且通常“通过西方价值、情境与关切来建构框架”。^[17]

我希望把注意力从这种带有精英主义色彩、制定各种意义上高层级AI原则的事务，转向世界不同地方为回应AI实际问题而出现的、更地方化的AI技术与AI伦理日常遭遇。我通过AI的一种具体应用来观察这一点：人脸识别摄像头和软件，尤其是其执法用途。这是地方层面介入AI并形成抵抗的一个具体例子；在包括英国在内的一些地方，这些过程催生了地方化的AI治理形式。尽管更全球性的AI伦理与治理倡议许诺崇高理想和大规模影响的潜力，尽管对AI及人脸识别等全球化、跨国技术和应用而言，更全球性的进路很可能更加适当，本章所考察的仍是人与AI技术及AI伦理在日常、地方层面的遭遇。迄今，大量AI伦理文献与活动都忽略、冷落了地方与日常，原因或许在于这些遭遇被认为“影响力”较小。然而，若不了解这些地方遭遇，高层级AI伦理就会保持抽象、悬浮，且往往去政治化。

无论如何，当个人与社区协商、质疑AI的某些用途，而政策制定者和法律可能回应他们的愿望、从而带来变化时，这些日常遭遇会以其他方式产生影响。这显然会影响事情发生的地方，却缺少承认，而且可能与传统高层级AI伦理倡议所包含的、无论隐含还是明确的普遍性主张形成反差；学术研究中的影响力要求，鼓励了这种普遍性主张。^[18]至少就人脸识别而言，或许在更广泛领域也是如此，对这一应用的

地方性与日常遭遇给予更多AI“伦理”关注，可以突显或服务于行动主义、抵抗或批评；而瞄准更抽象、更高或“普遍”层面的伦理关注，则经常更服务于资本和政治权力。^[19]

我首先考察AI已经作为日常技术存在的方式。我把人脸识别作为日常AI的实例：它的使用引发了争议，在一些地方导致使用受到限制；本文尤其关注英国。总体而言，这表明，与AI相遇的关键接点，也就是开展伦理、法律与政治追问的关键场所，正是、也必须是个人和社区与AI发生关系，并在某些情况下——例如人脸识别——对AI提出争议的地方。^[20]超越技术官僚式的高层级AI伦理规范制定，对这些日常遭遇的考量——包括抗议、社会运动和通过诉讼开展的法律动员——必须成为AI伦理讨论的一部分，特别是在英国这样的情境中：日常经验呈现的图景，与英国高层级AI战略和政策的想象截然不同。

作为日常技术的AI

AI正成为世界各地的日常技术，尽管人们往往不这样看待它。一些作者已探讨AI中的日常观念，以及人们关于AI的日常实践与经验，其中包括 Burgess、Mitchell 和 Highfield。他们的目标是：

超越当前围绕自动驾驶汽车、算法和机器人技术的炒作与焦虑，更精确、更扎实地理解自动化究竟可能意味着什么；它如何与日常生活纠缠在一起，产生什么效果；以及研究这些关系，如何也能帮助我们理解社会中更广泛的媒体变化过程。^[21]

此外，Pink 等人认识到：

对这些自动化技术的讨论，常常笼罩在突出极端、壮观案例的叙事中，而不是关注普通、平凡的现实；后者才构成绝大多数人实际遭遇这些技术时的特征。^[22]

AI正渗入我们的日常生活，尽管方式各异、情境不同；这种对日常的关注，与许多AI文献及其他讨论不同。^[23]后者专注于更全球或更抽象的层面，而且也往往发生在更精英化的层面，正如上文 Hagendorff 所指出的。人与AI的遭遇发生在日常之中，即使日常遭遇在不同场景下可能呈现不同面貌。

不过，日常也是人们能够反击技术——包括AI与自动化——的地方，尽管相关辩论和文献常常暗示人们只能被动承受。对 Pink 等人而言：

普通公民被描述为被动地受制于数字数据经济支持者的操纵与剥削。然而，自动化逻辑并非处处相同——它也并非在每一次使用、每一个地理位置，都以同样的强度运作。人们能够抵抗，而且确实在抵抗……^[24]

对许多人——即便不是大多数人——来说，与AI的遭遇主要发生在这种日常、地方化层面；同样，大量真正具有“约束力”的AI治理也发生在这一层面，而我认为，迄今许多AI辩论都忽视了它。这种治理能够由遭遇AI、与之协商、并在某些情况下抵抗AI的个人与社区来塑造，就像他们对待其他数据驱动监控技术一样。^[25]这正是我稍后通过人脸识别视角讨论的内容：AI伦理如何在具体、地方层面展开，以及它与“更高层级”的AI伦理讨论和表述有何关联，或没有关联。首先，我借助法律研究中的日常观念，考察日常AI法律与伦理意味着什么。

从AI伦理转向法律，再转向日常

对法律制定和规范形成的考量，需要纳入这种日常AI观念，因为在某些情况下，围绕规范的日常协商与争议，比近年我们看到的高层级、抽象化AI伦理活动，更有影响力或更令人满意地处理了AI伦理问题。

Wagner 将转向这类高层级AI伦理倡议的做法批评为“伦理洗白”，因为伦理声明和倡议在实施中通常缺少法律或其他形式的可执行性与问责。^[26]因此，它们成了具有约束力的权利与责任的替代品，而非补充。必须注意，在AI治理情境中，伦理是以一种特定方式使用的：通常由民族国家和大企业推广一系列无约束力的规范；对伦理的批评针对的是这种具体情境和用法。但伦理还有更广泛的意义，因为法律和其他规范体系也都是应用伦理的表现形式。^[27]

不过，仅有AI规范的法律可执行性，未必足够，也未必适合处理AI伦理原则不可执行的问题，因为这些规范的内容及其可执行性，都需要是“好的”。^[28]美国 Trump 政府采纳了具有法律约束力的AI行政命令，要求对这项技术采取放松监管的进路；批评无约束力AI伦理的人，大概既不会寻求、也不会满足于这种结果。^[29]无论如何，具有法律可执行性的AI伦理／治理倡议仍然很少，而已有倡议也不在国际层面，而是在区域或国家层面。

在国际层面，联合国教科文组织成员国最近通过了《人工智能伦理问题建议书》。这一举措意义重大，因为它是该主题首个全球标准；然而，它对签署国没有约束力，只是“建议”成员国在各自国内司法管辖范围内，以“自愿方式”实施。^[30]迄今，欧盟制定自身AI立法《欧盟人工智能法》的努力受到大量关注；本章撰写时，该法案仍在讨论之中。^[31]它值得关注，因为这是一个全球主要司法管辖区首次作出以有约束力方式监管AI的重要尝试，不过，按目前形式，它不会完全禁止执法部门使用人脸识别。^[32]

但在这里，我想考察的，是地方或微观层面——而非国家或国际层面——对AI伦理规范及法律更日常的理解、协商和抵抗。这样做，是为了与“日常法律”或“法律社会化”的研究对话：它们探讨人们如何体验、形成并遵守（法律）规范，^[33]或者用 Sarat 和 Kearns 的话说，“法律的消费者如何生产他们自己的法律，并在这

一过程中转化和再生产国家法律”。^[34]这是因为，来自个人与社区、针对AI使用——尤其是国家与企业使用——的这些理解、协商与抵抗，能让我们认识到，人们注意到了哪些AI用途，又觉得哪些可以接受、哪些不可接受；这反过来可能影响国家法律和企业实践。人脸识别技术值得注意，因为它在世界不同地方、不同情境中的使用引发了实体抗议，而且在英国成为诉讼与政策变更的基础。

在这一点上，我还想把关于日常法律的讨论，与法律如何同社会运动和抗议相互作用联系起来；这是社会运动学者与法律学者都研究不足的领域。^[35]这对于人脸识别十分重要，因为抗议和倡议行动积累了压力，导致这项实践遭到禁止或暂停，也通过诉讼对其使用提出挑战。这还与 Milan 等人关于“数据行动主义”的研究相联系：这种行动主义“批判性地介入数据对社会生活的多重影响”，并包括“例如，对与大数据相伴而生的歧视、社会排斥和隐私侵犯作出反霸权回应的社会技术实践”。^[36]数据行动主义尤其强调“由普通人、非政府组织和社会运动网络共同表达的[针对数据化的]草根争议过程”。^[37]无论以社会运动回应，还是以立法和政策回应的形式出现，对人脸识别的反对都构成 Kazansky 所谓“对数据驱动监控的抵抗”。^[38]然而，在这一领域关于新技术——尤其AI——的文献中，抗议、社会运动与法律/政策变化很少被放在一起考察。

我引入这些概念，作为考察人脸识别的背景：它既是一种悄然进入世界各地人们生活的日常AI技术，也是一处社会运动和数据行动主义争议的场所，与AI的法律和伦理相互作用。我们有必要对AI、行动主义和伦理——包括法律——开展更多理论与实证研究，以获得更深的理解，特别要从日常视角理解普通人如何遭遇并介入这些问题。我在这里力图引入这些议题，但还可以直接开展更多工作，例如与那些从日常视角影响、协商、尤其是抵抗人脸识别，而通常不参与“更高层级”AI伦理倡议和规范形成的人一起研究。

日常AI法律、伦理与抗议已经是实践现实。一个例子是，2020年英格兰学生毕业成绩由算法决定——传统考试因COVID-19疫情取消——结果出现不公，引发针对教育部的示威；年轻人高喊并举着写有“去他妈的算法”（Fuck the Algorithm）的标语。Kaun 将其视为 Willim 所说的数字技术“日常化”的例子，即“对影响我们日常生活的复杂技术形成日常理解”。^[39]公共部门使用算法，也引发了更广泛的争议，例如澳大利亚的 RoboDebt 福利监控丑闻。^[40]2020—2021年印度农民抗议中，也可见日常AI伦理的其他例子：农民理解了相关计划之间的联系，综合企业 Jio——其众多业务包括运营移动通信网络——计划进入农业科技领域，并利用面向农民的AI交易平台巩固自身权力；许多农民因此抵制该运营商，把移动通信服务转到竞争者那里。^[41]

这些例子表明，围绕AI的争议已经存在于人们的日常生活中，并激发地方性行动，包括抗议；这些行动可能推动法律和政策变化。人与AI及其政治的这些日常遭遇，把AI伦理从遥远的政策制定者以及政治、企业精英手中，带（回）个人与社区，承

认他们／我们在协商和抵抗技术应用方面的能动性。这些争议与抵抗能够通过激发行动和变革，在具体地方层面限制AI用途，从而处理Wagner以“伦理洗白”所批评的执行力缺口。相比之下，那些高远、时而精英主义的AI伦理倡议，往往缺乏“约束力”，通常也不会禁止或严厉限制AI的某些用途与应用。

作为日常AI的人脸识别

在这里，我希望聚焦于以人脸识别形式出现的AI应用，以及世界不同地方的人们与它的日常遭遇；其中一些遭遇催生了日常AI法律与伦理。我重点讨论英国的人脸识别经验，因为这是我最熟悉的地理区域，而且我们在这里经历了与日常人脸识别使用相关的抗议、政策和法律事件，英国不同地区对人脸识别使用的进路也有所不同。这些都可以与英国高层级AI政策中“促进创新”和新帝国主义的取向并置考察。

人脸识别是一种从数字图像中识别个人的技术，通常把此人的面部特征与数据库中储存的面部生物识别图像进行比较。当图像捕捉和分析借助AI实时进行时，例如公共场所中的“智能”闭路电视摄像头，人脸识别可以是“实时”的。人脸识别因其不准确性而备受争议，尤其是识别女性比识别男性、识别有色人种比识别白人更容易出错，“肤色较深的女性是被误分类最多的群体”。^[42]此外，人脸识别技术研究、开发和试验的条件也引发争议：例如，中国的人脸识别产品被用于针对新疆／东突厥斯坦的维吾尔人及其他少数民族裔；^[43]西方的Clearview AI则在用户不知情或未经许可的情况下，从社交媒体抓取照片，其产品被美国执法部门、也可能被欧洲执法部门使用。^[44]最近，Clearview的这些抓取行为在欧盟、英国和澳大利亚引来了数据保护侵权认定与罚款。^[45]

人脸识别已在全球各种社会、政治与经济情境中实施，既存在于威权政体，也存在于（所谓的）自由民主国家。因此，它正成为一种日常AI技术，公众在处理日常事务时，尤其在公共场所，会遭遇它。关键在于，在某些情况下，这些与人脸识别的日常遭遇，引出了公众的协商过程和直接抵抗。人脸识别已成为社会运动动员的对象，有的是专门反对这种监控技术的使用，有的则是更广泛抗议的一部分。日常法律也被动员起来反对人脸识别，进而引出国家法律应如何处理它的问题。

人脸识别和闭路电视摄像头，已在全球多个地点成为抗议乃至实际破坏的对象。2019年伊朗反对政府提高汽油价格的抗议期间，^[46]出现了抗议者在该国不同地方，包括设拉子和德黑兰，使闭路电视摄像头失效并将其毁坏的影像。^[47]在较近的2021年胡齐斯坦省抗议中，也有影像似乎显示了针对闭路电视摄像头的类似袭击。^[48]伊朗存在广泛的监控基础设施，尤其在2019年抗议之后。据Akbari所述，此后“咖啡馆、大学，甚至幼儿园都被要求安装闭路电视摄像头；大城市的交通控制摄像头大量涌现”，而且“政府积极使用闭路电视／交通摄像头影像来打击政治异议”。^[49]

同样在2019年，在“对技术的不信任持续存在”的香港，疑似具有人脸识别功能的闭路电视摄像头，成为反对引渡法案和国家安全法的抗议者的目标。^[50]不仅“口罩、雨伞和激光……被示威者经常用于遮蔽闭路电视摄像头……从而使人脸识别失效”，^[51]抗议者还在一场反对监控及人脸识别技术日益普及的抗议中，拆倒了政府安装的新型“智能”灯柱；这些灯柱的全部技术能力并未公开。^[52]

抗议者不仅拆倒灯柱，还“解剖”它们，打开其“黑箱”，查看里面究竟有哪些组件和设备，包括是否含有人脸识别设备，因为香港政府此前声称这些灯柱仅用于监测空气质量与交通。^[53]一些智能灯柱内确实有摄像头；虽然这些摄像头似乎没有人脸识别能力，独立专家却认为，修改摄像头、加入这类功能并不困难。^[54]无论如何，出于隐私方面的担忧，香港当局决定不启用智能灯柱的某些功能。

在伊朗与香港的两个例子中，公共场所中人脸识别技术的可能存在或实际存在，促成了抗议与动员。它们可以被理解为回应物质处境、反抗国家权力的更广泛运动的一部分。不过，两场运动中都值得注意的，是公众对（可能具有）人脸识别功能的闭路电视摄像头的怀疑，以及对其采取实体行动；这显示了公民抵抗数字数据（政治）经济某些方面的形式。在香港，这促使当局决定不实施智能灯柱的某些功能；在《国家安全法》的背景下，这是对抗议者关切一次值得注意、也颇为罕见的积极回应，同时也表明，在香港这类自上而下智慧城市项目的总体背景中，政府能够回应公民的关切。^[55]

英国的人脸识别、日常AI法律与伦理

在英国，作为日常AI的人脸识别，以及围绕它的争议，一直十分突出；反对人脸识别的动员导致了诉讼与政策变化，也使英国不同地区的进路出现分化。实时人脸识别技术已在英国不同地方被用于公共场所执法，引起日益强烈的争议与法律挑战。因此，我认为它是一个有意思的日常AI伦理及法律案例，也能展示：个人与社区在更地方化层面遭遇、协商和抵抗AI的活动，如何对更广泛的AI治理产生影响。此外，英国内部对待人脸识别的不同进路，也表明除国家、大洲及国际层面之外，还必须关注地方层面。如前所述，英国在“高”层面对AI采取促进创新的技术解决主义进路；但只有考察人脸识别与公众之间这些更地方化的遭遇，AI的“阴暗面”以及围绕它的民主争议才会清楚呈现。在美国，同样能看到围绕人脸识别的争议引发法律和政策变化；一些市政当局已经禁止警方使用人脸识别，包括率先这样做的旧金山。^[56]弗吉尼亚州和佛蒙特州也禁止警方在全州境内使用人脸识别。^[57]地方层面针对实时人脸识别等有问题AI用途的动员，可以带来禁令；这些禁令又能像滚雪球一样流传，启发其他地方出台禁令，围绕AI形成自下而上、更具批判性的规范，有别于自上而下却往往缺乏约束力的AI伦理倡议。

近来在英国，已有一段在争议性情境下提议使用、或实际使用人脸识别技术的历史，尤其涉及警方与执法部门；即使最近暂停这些用途的苏格兰，也不例外。例

如，早在2016年，苏格兰职业足球联赛（SPFL）就打算在苏格兰足球场引入人脸识别技术；当时，利用摄像头监控球迷的做法正在加强，球迷与警方的关系不断恶化。^[58]多个球迷支持者组织公开反对这些计划，包括在比赛中展开反人脸识别横幅。^[59]苏格兰警方也表明，希望在更广泛的活动中使用实时人脸识别，而不只针对球迷。苏格兰最大城市格拉斯哥于2015年为市中心购入了具有人脸识别功能的摄像头，但出于隐私与人权担忧，这些摄像头一直未被使用。^[60]在这两个案例中，那些将成为技术使用对象的人提出的压力与关切，促使公共机构重新考虑并克制使用人脸识别；不过，在后一个案例中，购买从未使用的技术，也造成了公共资金浪费。这颇具讽刺意味，因为在紧缩和私有化背景下，警方受到鼓励，转向人脸识别供应商等私人科技服务商，而这样做据称是为了削减成本。^[61]

2020年，苏格兰议会司法小组委员会审议了警方使用人脸识别技术的问题。委员会开展了咨询程序，我和其他多位学者、民间社会团体都参与其中；我们大多数人质疑警方使用人脸识别，指出其歧视性特征，以及其他地方——尤其美国城市——的禁令。委员会得出结论：“苏格兰警方没有正当依据投资这项技术”，主要原因是该技术涉及性别和种族歧视。^[62]自此，苏格兰警方使用实时人脸识别技术一直处于暂停状态。

这与英格兰和威尔士采取的进路形成对比：英格兰和威尔士属于同一个司法管辖区，苏格兰与北爱尔兰分别构成英国的另外两个司法管辖区。据 Big Brother Watch 所述：

英国警察机关推出自动人脸识别的速度，超过世界任何其他民主国家。莱斯特郡警方、南威尔士警方和伦敦大都会警察局，已在购物中心、节庆、体育赛事、音乐会、社区活动，甚至和平示威中部署这一技术。有一支警察队伍甚至利用这种监控工具，让患有精神健康问题的无辜者远离一场公共活动。^[63]

在伦敦，大都会警察局曾在人口中黑人与少数族裔（BAME）比例较高的地区和使用人脸识别，例如东伦敦斯特拉特福，以及2016年和2017年的诺丁山狂欢节，尽管人脸识别对黑人的识别存在不准确问题。这样做也无视了伦敦大都会警方与黑人社区之间本已紧张的关系。^[64]一些具体的监控和数据收集活动，对黑人社区产生了不成比例的影响，包括为大都会警察局颇具争议的“帮派矩阵”（Gangs Matrix）数据库收集涉嫌参与帮派活动者的信息，而其中大多数人是黑人。^[65]英国数据保护机构信息专员办公室（ICO）认定，该矩阵不符合数据保护法；但大都会警方接到的命令不是销毁它，而是使它符合这些规范。该矩阵依然引发争议，本章撰写时，正在面对由公民自由与人权非政府组织 Liberty 牵头的另一项法律挑战；此次理由包括违反种族歧视法，以及人权、数据保护和公法原则。^[66]

与伊朗和香港情境不同，英国的人脸识别摄像头没有遭受实体袭击，但当它们用于日常公共场所时，仍引起至少一部分公众的强烈本能反应。一些人得知公共场所部

署了人脸识别摄像头后，遮住面孔以保护隐私；至少有一人因此被警方罚款。^[67] 苏格兰球迷也为此展开专门反对在球场使用人脸识别的横幅。这些可被视为对英国警察“过度军事化”的更广泛争议的一部分；Weitzberg 将人脸识别的使用列为这一趋势的组成部分。^[68] 更广泛的公众对人脸识别的支持有限，甚至警方内部某些人士也持怀疑态度。Ada Lovelace Institute 关于公众对人脸识别态度的全国调查显示，大多数公众希望政府限制警方使用人脸识别，并反对将该技术用于商业用途。^[69] Urquhart 和 Miranda 对英国一线警员的研究也显示，即便警员的立场也“主要是对这项技术持怀疑和不信任态度”。^[70]

英国实时人脸识别的批评者，也通过诉讼动员法律，尤其是人权法与数据保护法，取得了“世界任何地方首次对警方使用自动人脸识别技术提出重大且成功的法律挑战”。^[71] 公民自由活动者 Ed Bridges 挑战南威尔士警方使用实时人脸识别，理由是这侵犯了隐私权，并违反数据保护法和平等法。一审中，高等法院认定，虽然人脸识别确实干涉公众权利，但由于规范人脸识别使用的框架中存在保障措施，南威尔士警方的使用是合法的。^[72] 然而，这项判决在上诉中被推翻：上诉法院认定，实时人脸识别的使用确实侵犯了人权，其治理框架存在“根本缺陷”，而且警方没有确保所用软件在种族和性别方面不带偏见。^[73] 南威尔士警方不会对这一判决提出上诉。不过，与前述ICO关于大都会警方“帮派矩阵”的决定一样，上诉法院并未认定警方在公共场所使用人脸识别本身违法，而只是认为缺乏适当保障；事实上，“该判决仍然肯定了自动人脸识别在现代警务和执法中的作用”。^[74]

综上，苏格兰实际上禁止警方使用实时人脸识别，而英格兰和威尔士则在一些限制下允许使用；对于英国警方把实时人脸识别作为公共场所的日常AI应用而言，这形成了 Lynch 所说的“南北分野”：

例如，如果你走在伦敦或威尔士某些地方的街头，实时人脸识别技术现在能够在未经同意的情况下扫描你的脸，你甚至可能接受当场身份检查，尤其如果你是女性或少数族裔。但在苏格兰，你不必担心这一点——至少目前如此。^[75]

不过，警方使用人脸识别，只是图景的一部分。英国日常生活中还有其他引发争议的人脸识别用途，包括苏格兰。2021年，北艾尔郡九所学校使用了人脸识别技术，以便食堂午餐更快、无接触地付款。^[76] ICO敦促地方政府采用侵入性较低的方式，以确保满足必要性与相称性要求；不久之后，人脸识别的使用似乎就被暂停。^[77] 这或许把苏格兰学校的人脸识别扼杀在萌芽阶段，但传统上被视为有道德零售商的合作社超市 Co-op，却正以安全和安保为由，在英格兰南部的门店使用实时人脸识别。数字权利团体 Big Brother Watch 反对这一做法，并发起了 #StopCoopSpying 社交媒体运动。^[78] 本章撰写时，Big Brother Watch 与数字权利机构 AWO 正向ICO投诉 Co-op。^[79]

英国人与人脸识别的日常遭遇及争议表明，从警方在公共场所部署，到学校和超市中的使用，公众如何在平凡生活中以人脸识别的形式接触AI。对政策制定者最成功的影响，体现在苏格兰议会暂停警方使用实时人脸识别的做法。法律挑战和利用ICO投诉程序的行动——尤其是活动人士与非政府组织所为——在约束人脸识别方面取得了一些成效，但并非彻底胜利。ICO或法院不愿直接认定在公共场所使用侵入性极强的人脸识别违法，说明自下而上的规范形成只取得了部分成功；不过，这也可能反映这些机构对英国议会的尊让，因为它们或许认为议会才是实施此类禁令更合适的机构。然而，我们仍在等待这样的行动，尽管最近 Matthew Ryder 王室御用大律师关于英格兰及威尔士生物识别数据治理的独立审查，加强了这类呼声。他建议，需要为生物识别技术的所有用途建立新的立法框架，并分别为警方及其他实时人脸识别使用者制定具有法律约束力的操作守则；在这些措施实施之前，一切实时人脸识别使用都应停止。^[80]

在地方情境中形成的规范，可以在更广的国际范围流通。有时，这源于国家或全球资本的流动，例如美国加利福尼亚州形成的法律与政策；欧盟法律与政策也日益在全球产生影响，其治理机制的“布鲁塞尔效应”影响其他地方的法律和政策。这是因为欧盟是世界最大的贸易集团，并积极发展和传播其法律与监管，使之超出自身边界。^[81]欧盟拟议中的《人工智能法》也可能如此，追随《通用数据保护条例》（GDPR）成为事实上的全球规范；^[82]但如上所述，该法目前不会彻底禁止人脸识别的使用，即便是执法用途。

不过，苏格兰的例子表明，还可能存在并不服务于全球资本与权力的其他规范流通形式：苏格兰议会司法小组委员会呼吁暂停苏格兰警方使用实时人脸识别时，援引了世界其他地方对警方使用人脸识别的地方性禁令。这表明，通过协商与质疑AI用途而在地方形成的规范，也能够以更广的全球范围流通，并影响其他地方的活动。由此，地方化AI规范可能形成滚雪球效应，被其他地理情境中的社会运动、抗议及法律动员所运用。

在“高层级”上，英国阐明了其AI公共研究资助进路，以及针对包括人脸识别在内的AI采取“轻触式”、无约束力治理的政策意图。这进一步显示出英国与欧盟监管AI意图之间的分歧。它可被视为英国脱欧后轨迹的一部分；这一轨迹还包括疏远欧盟的数据保护制度，也符合 Ossewaarde 和 Gulenc 前述观察：英国AI进路具有数字乌托邦、技术解决主义和新帝国主义色彩。^[83]个人与社区在英国公共空间中的日常遭遇，以及对实际人脸识别使用的争议，表明这些逻辑如何延续，也如何遭到抵抗，尤其当人脸识别成为英国过度军事化执法的一部分、针对黑人与少数族裔社区时。在某些情况下，当机会出现，这些动员可以引发诉讼——尽管迄今成功有限——及地方政策变化。英国研究者、活动人士和其他人，可能发现影响英国中央政府的前景有限，但人脸识别这一案例表明，可以通过诉讼施加压力。影响更地方化的治理结构，例如苏格兰、威尔士和北爱尔兰的权力下放行政机构，可能有更多机

会，也可能更有希望产生影响。目前，执掌这些行政机构的政党，均不同于在英国议会掌权的保守党；出于政治原因，它们可能希望让政策与英国政府有所区别，在脱欧和COVID-19情境下，这种意愿更加突出，从而导致政策与治理的碎片化和差异化。^[84]此外，全球资本、尤其大型科技企业，可能较少尝试以有利于自身利益的方式影响这些行政机构。这些条件为地方层面协商和抵抗AI提供了可能性，在约束人脸识别方面，这些可能性已在某种程度上实现；它们可以与英国政府对AI放任式、“促进创新”的进路相对照。

结论

此时此刻，在世界各地的地方与日常层面，AI及AI伦理的实践和应用正在发生，普通公众以无数形式遭遇这项技术。然而，AI伦理讨论与倡议很少聚焦这些遭遇，以及随之而来的协商与争议。通过英国人脸识别的案例，我展示了关注地方如何成为理解AI伦理在实践中怎样展开、怎样形成、又受到什么塑造的关键；这种实践有时能带来真正具有“约束力”的法律与政策变化，服务于个人与社区，而非国家权力与资本。这种“约束力”，正是经过伦理洗白的高层级AI伦理倡议往往缺乏的。因此，我们需要更多参与地方社会运动以及日常法律与政策，努力建设、形成并塑造更好的AI。这正是真实变革能够在当下发生的地方，而这种变革不一定服务于政治与经济权力。

致谢

感谢 Xaroula Kerasidou、评阅人 Jake Goldenfein 与 Nancy Salem，以及编辑 Thao Phan 对本文早期版本提出意见。

资助披露

本研究没有特定资金来源资助，是在我的受薪工作时间及这些时间之外开展的。撰写本文期间，我持有或曾持有UKRI资助的项目经费，但它们并未支持本文。

参考文献

- Ada Lovelace Institute. ‘Beyond Face Value: Public Attitudes to Facial Recognition technology’, 2019, <https://www.adalovelaceinstitute.org/report/beyond-face-value-public-attitudes-to-facial-recognition-technology/>. [中文题名：不止于表面：公众对人脸识别技术的态度。]
- Akbari, Azadeh. ‘The Threat of Automating Control: Surveillance of Women’s Clothing in Iran’, in Aleš Zavrñnik and Vasja Badalič (eds) *Automating Crime Prevention, Surveillance, and Military Operations*, Cham: Springer, 2021, pp. 183–99. [中文题名：自动化控制的威胁：伊朗对女性服装的监控；载《犯罪预防、监控与军事行动的自动化》。]

- Beraldo, Davide and Milan, Stefania. 'From Data Politics to the Contentious Politics of Data', *Big Data & Society* 6.2 (2019). [中文题名: 从数据政治到数据的争议政治。]
- Bietti, Elettra. 'From Ethics Washing to Ethics Bashing: A View on Tech Ethics from Within Moral Philosophy', *Proceedings of ACM FAT* Conference*, 2020, <https://ssrn.com/abstract=3513182>. [中文题名: 从伦理洗白到伦理抨击: 道德哲学内部的科技伦理观。]
- Big Brother Watch. 'Face Off: The Lawless Growth of Facial Recognition in UK Policing', 2018, <http://bigbrotherwatch.org.uk/wp-content/uploads/2018/05/Face-Off-final-digital-1>. [中文题名: 面孔对抗: 英国警务中人脸识别不受法律约束的扩张。]
- Bradford, Anu. *The Brussels Effect: How the European Union Rules the World*. Oxford: Oxford University Press, 2020. [中文题名: 布鲁塞尔效应: 欧盟如何统治世界。]
- Buolamwini, Joy and Gebru Timnit. 'Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification' in *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, PMLR 81, 2018, 77–91. [中文题名: 性别色差: 商业性别分类中交叉性的准确率差异。]
- Burgess, Jean, Mitchell, Peta, and Highfield, Tim. 'Automating the Digital Everyday: An Introduction', *Media International Australia*, 166.1 (2018): 6–10. [中文题名: 日常数字生活的自动化: 导论。]
- Chubb, Jennifer and Reed, Mark. 'The Politics of Research Impact: Academic Perceptions of the Implications for Research Funding, Motivation and Quality', *British Politics* 13 (2018): 295–311. [中文题名: 研究影响的政治: 学者对其在研究资助、动机和质量方面影响的想法。]
- Daly, Angela. 'Algorithmic Oppression with Chinese Characteristics: AI Against Xinjiang's Uyghurs' in *Artificial Intelligence: Human Rights, Social Justice and Development: Global Information Society Watch 2019 Report*, 2019, pp. 108–112. [中文题名: 具有中国特色的算法压迫: 针对新疆维吾尔人的AI; 载《人工智能: 人权、社会正义与发展——2019年全球信息社会观察报告》。]
- Daly, Angela, Devitt, S. Kate and Mann, Monique. 'AI Ethics Needs Good Data' in Pieter Verdegem (ed) *AI for Everyone? Critical Perspectives*, London: University of Westminster Press, 2021, pp. 103–22. [中文题名: AI伦理需要良好数据; 载《人人共享AI? 批判性视角》。]
- Daly, Angela, Hagendorff, Thilo, Hui, Li, Mann, Monique, Marda, Vidushi, Wagner, Ben, and Wang, Wayne Wei. 'AI, Governance and Ethics: Global Perspectives' in Hans Micklitz, Oreste Pollicino, Amnon Reichman, Andrea Simoncini, Giovanni Sartor and Giovanni De Gregorio (eds) *Constitutional Challenges in the Algorithmic Society*, Cambridge: Cambridge University Press, 2022, pp. 182–201. [中文题名: AI、治理与伦理: 全球视角; 载《算法社会中的宪制挑战》。]
- Daly, Angela. 'Neo-Liberal Business-As-Usual or Post-Surveillance Capitalism with European Characteristics? The EU's General Data Protection Regulation in a Multi-Polar Internet' in Rolien Hoyng and Gladys Pak Lei Chong (eds), *Critiquing Communication Innovation: New Media in a Multipolar World*, East Lansing: Michigan State University Press, 2022 (forthcoming). [中文题名: 新自由主义照常运转, 还是具有

- 欧洲特色的后监控资本主义？多极互联网中的欧盟《通用数据保护条例》；载《批判传播创新：多极世界中的新媒体》（即将出版）。]
- Davies, Pascale. ‘UK Schools Suspend Use of Controversial Facial Recognition Technology’, *Euronews*, 25 October 2021, <https://www.euronews.com/next/2021/10/18/schools-in-scotland-start-using-facial-recognition-on-children-paying-for-lunch>. [中文题名：英国学校暂停使用引发争议的人脸识别技术。]
- Dearden, Lizzie. ‘Police Stop People for Covering Their Faces From Facial Recognition Camera then Fine Man £90 After he Protested’, *The Independent*, 31 January 2019, <https://www.independent.co.uk/news/uk/crime/facial-recognition-cameras-technology-london-trial-met-police-face-cover-man-fined-a8756936.html>. [中文题名：警方拦下为躲避人脸识别摄像头而遮脸的人，一名男子抗议后被罚90英镑。]
- Edwards, Rosalind. ‘Why Do Academics Do Unfunded Research? Resistance, Compliance and Identity in the UK Neo-liberal University’, *Studies in Higher Education* 47.4 (2022): 904–14. [中文题名：学者为什么开展无资助研究？英国新自由主义大学中的抵抗、服从与身份。]
- Elliott, Ian, Bottom, Karin, Carmichael, Paul, Liddle, Joyce, Martin, Steve, and Pyper, Robert. ‘The Fragmentation of Public Administration: Differentiated and Decentered Governance in the (dis)United Kingdom’, *Public Administration* 100.1 (2022): 98–115. [中文题名：公共行政的碎片化：（不）联合王国中的差异化与去中心化治理。]
- Ewick, Patricia and Silbey, Susan. *The Common Place of Law: Stories from Everyday Life*, Chicago: University of Chicago Press, 1998. [中文题名：法律的日常场所：日常生活的故事。]
- Feathers, Todd. ‘Facial Recognition Is Racist. Why Aren’t More Cities Banning It?’, *Vice*, 25 May 2021, <https://www.vice.com/en/article/4avx3m/facial-recognition-is-racist-why-arent-more-cities-banning-it>. [中文题名：人脸识别具有种族歧视性，为什么没有更多城市禁止它？。]
- Finn, Mike. *British Universities in the Brexit Moment: Political, Economic and Cultural Implications*. Bingley: Emerald Publishing, 2018. [中文题名：脱欧时刻的英国大学：政治、经济与文化含义。]
- Gleeson, Sean. ‘How Smart are Hong Kong’s Lampposts?’, *AFP Fact Check*, 4 September 2019, <https://factcheck.afp.com/how-smart-are-hong-kongs-lampposts>. [中文题名：香港的灯柱有多智能？。]
- Hagendorff, Thilo. ‘The Ethics of AI Ethics: An Evaluation of Guidelines’, *Minds & Machines* 30 (2021): 99–120. [中文题名：AI伦理的伦理：指南评估。]
- Hagendorff, Thilo. ‘Blind Spots in AI ethics’, *AI Ethics* (2021). [中文题名：AI伦理的盲点。]
- Hamed, Assiah. ‘Co-op Defends Facial Recognition Cameras in Bristol Stores Amid Claims of ‘Orwellian’ Surveillance’, *Bristol Post*, 9 December 2021, <https://www.bristolpost.co.uk/whats-on/shopping/co-op-defends-facial-recognition-6302476> (accessed 1 February 2022). [中文题名：被指实施“奥威尔式”监控之际，Co-op为布里斯托门店的人脸识别摄像头辩护。]

- Hamilton-Smith, Niall, McBride, Maureen, and Atkinson, Colin. 'Lights, Camera, Provocation? Exploring Experiences of Surveillance in the Policing of Scottish Football', *Policing and Society* 31.2 (2021): 179–94. [中文题名: 灯光、摄像、挑衅? 探究苏格兰足球警务中的监控体验。]
- Heikkilä, Melissa. 'The Walls are Closing in on Clearview AI', *MIT Technology Review*, 24 May 2022, <https://www.technologyreview.com/2022/05/24/1052653/clearview-ai-data-privacy-uk/>. [中文题名: Clearview AI四面楚歌。]
- Hunton Andrews Kurth. 'UK Court of Appeal Finds Automated Facial Recognition Technology Unlawful in Bridges v South Wales Police', 12 August 2020, <https://www.huntonprivacyblog.com/2020/08/12/uk-court-of-appeal-finds-automated-facial-recognition-technology-unlawful-in-bridges-v-south-wales-police/>. [中文题名: 英国上诉法院在Bridges诉南威尔士警方案中认定自动人脸识别技术违法。]
- Jobin, Anna, Ienca, Marcello, and Vayena, Effy. 'The Global Landscape of AI Ethics Guidelines', *Nature Machine Intelligence* 1 (2019): 389–99. [中文题名: AI伦理指南的全球格局。]
- Kaun, Anne. 'Suing the Algorithm: the Mundanization of Automated Decision-making in Public Services Through Litigation', *Information, Communication & Society* (2021). [中文题名: 起诉算法: 通过诉讼使公共服务中的自动化决策日常化。]
- Kazansky, Becky, Torres, Guillen, van der Velden, Lonneke, Wissenbach, Kersti, and Milan, Stefania. 'Data for the Social Good: Towards a Data-Activist Research Agenda' in Angela Daly, S. Kate Devitt and Monique Mann (eds) *Good Data*, Amsterdam: Institute of Network Cultures, 2019, pp. 244–59. [中文题名: 服务社会公益的数据: 迈向数据行动主义研究议程; 载《良好数据》。]
- Kazansky, Becky. '“It Depends on Your Threat Model”: The Anticipatory Dimensions of Resistance to Data-driven Surveillance', *Big Data & Society* 8.1 (2021). [中文题名: “取决于你的威胁模型”: 抵抗数据驱动监控的预期性维度。]
- Kazim, Emre, Almeida, Denise, Kingsman, Nigel, Kerrigan, Charles, Koshiyama, Adriano, Lomas, Elizabeth and Hilliard, Airlie. 'Innovation and opportunity: review of the UK's national AI strategy', *Discover Artificial Intelligence* 1.14 (2021): 1–10. [中文题名: 创新与机遇: 评述英国国家AI战略。]
- Kewalramani, Manoj and Seth, Rohan. 'Networked Protests & State Responses: The Case of Hong Kong 2019-2020', Takshashila Discussion Document 2020-03, 2020, <https://ssrn.com/abstract=3580591>. [中文题名: 网络化抗议与国家回应: 2019—2020年的香港案例。]
- Lee, Dave. 'San Francisco is First US City to Ban Facial Recognition', *BBC*, 15 May 2019, <https://www.bbc.co.uk/news/technology-48276660>. [中文题名: 旧金山成为美国首个禁止人脸识别的城市。]
- Leung, Kevin and Lee, H.Y. 'Implementing the Smart City: Who Has a Say? Some Insights from Hong Kong', *International Journal of Urban Sciences*(2021): 1–25. [中文题名: 实施智慧城市: 谁有发言权? 来自香港的一些观察。]

- Long, Lisa and Joseph-Salisbury, Remi. 'Black Mixed-race Men' s Perceptions and Experiences of the Police' , *Ethnic and Racial Studies* 42.2 (2019): 198–215. [中文题名: 黑人混血男性对警方的看法与经历。]
- Lu, Alex Jiahong. 'Toward Everyday Negotiation and Resistance Under Data-Driven Surveillance' , *Interactions* 29.2 (2022): 34–8. [中文题名: 迈向数据驱动监控下的日常协商与抵抗。]
- Lynch, Euan. 'The Use of Live Facial Recognition Technology in Scotland: A New North-South Divide?' , *UK Human Rights Blog*, 25 February 2020, <https://ukhumanrightsblog.com/2020/02/25/the-use-of-live-facial-recognition-technology-in-scotland-a-new-north-south-divide/>. [中文题名: 苏格兰实时人脸识别技术的使用: 新的南北分野?。]
- Mann, Monique. 'Social (In)security and Social (In)justice: Automation in the Australian Welfare System' in *Artificial intelligence: Human rights, social justice and development: Global Information Society Watch 2019 Report*, 2019, pp. 68–72. [中文题名: 社会(不)安全与社会(不)正义: 澳大利亚福利系统中的自动化; 载《人工智能: 人权、社会正义与发展——2019年全球信息社会观察报告》。]
- McCann, Michael. 'Law and Social Movements: Contemporary Perspectives' , *Annual Review of Law and Social Science* 2.1 (2006): 17–38. [中文题名: 法律与社会运动: 当代视角。]
- McGowran, Leigh. 'The Issues with the EU' s Draft Regulation on Facial Recognition AI' , *Silicon Republic*, 17 May 2022, <https://www.siliconrepublic.com/enterprise/the-issues-with-the-eus-draft-regulation-on-facial-recognition-ai>. [中文题名: 欧盟关于人脸识别AI的法规草案存在的问题。]
- Mega, Marcello. 'Cops Fear Gangsters are Evading Law as Glasgow' s Facial Recognition Cameras Remain Mothballed' , *Daily Record*, 11 August 2020, <https://www.dailyrecord.co.uk/news/scottish-news/scots-cops-fear-gangsters-evading-22499468>. [中文题名: 格拉斯哥人脸识别摄像头仍被封存, 警察担忧黑帮分子逃避法律。]
- Moule, Richard, Burruss, George, Gifford, Faith, Parry, Megan, and Fox, Bryanna. 'Legal Socialization and Subcultural Norms: Examining Linkages Between Perceptions of Procedural Justice, Legal Cynicism, and the Code of the Street' , *Journal of Criminal Justice* 61(2019): 26–39. [中文题名: 法律社会化与亚文化规范: 考察程序正义感知、法律犬儒主义和街头守则之间的联系。]
- Nijjar, Jasbinder. 'Police–school Partnerships and the War on Black Youth' , *Critical Social Policy*, 41.3 (2021): 491–501. [中文题名: 警校合作与针对黑人青年的战争。]
- ÓhÉigeartaigh, Seán, Whittlestone, Jess, Liu, Yang, Zeng, Yi, and Liu, Zhe. 'Overcoming Barriers to Cross-cultural Cooperation in AI Ethics and Governance' , *Philosophy and Technology* 33 (2020): 571–93. [中文题名: 克服AI伦理与治理跨文化合作的障碍。]
- Ossewaarde, Marinus and Gulenc, Erdener. 'National Varieties of Artificial Intelligence Discourses: Myth, Utopianism, and Solutionism in West European Policy Expectations' , *Computer* 53.11 (2020): 53–61. [中文题名: 人工智能话语的国家差异: 西欧政策预期中的神话、乌托邦主义与解决主义。]

- Parida, Tulsi and Ashok, Aparna. 'Consolidating Power in the Name of Progress: Techno-solutionism and Farmer Protests in India' in Frederike Kaltheuner (ed), *Fake AI*, Manchester: Meatspace Press, 2021, pp. 161–9. [中文题名: 以进步之名巩固权力: 印度的技术解决主义与农民抗议; 载《虚假AI》。]
- Phan, Thao, Goldenfein, Jake, Mann, Monique, and Kuch, Declan. 'Economies of Virtue: The Circulation of "Ethics" in Big Tech', *Science as Culture*, 31.1 (2022): 121–35. [中文题名: 美德经济: 大型科技企业中“伦理”的流通。]
- Pink, Sarah, Ruckenstein, Minna, Berg, Martin and Lupton, Deborah. 'Everyday Automation: Setting a Research Agenda' in Pink, Sarah, Berg, Martin, Lupton, Deborah, and Ruckenstein, Minna (eds) *Everyday Automation: Experiencing and Anticipating Emerging Technologies*, London: Routledge, 2022, pp. 1–20. [中文题名: 日常自动化: 设定研究议程; 载《日常自动化: 体验并预期新兴技术》。]
- Rezende, Isadora Neroni. 'Facial Recognition in Police Hands: Assessing the 'Clearview Case' From a European Perspective', *New Journal of European Criminal Law* 11.3 (2020): 375–89. [中文题名: 警方手中的人脸识别: 从欧洲视角评估“Clearview案”。]
- Ruthven, Graham. 'The Criminalization of Scottish Soccer Fans', *Vice*, 23 February 2016, <https://www.vice.com/en/article/9apyad/the-criminalization-of-scottish-soccer-fans>. [中文题名: 苏格兰球迷的犯罪化。]
- Ryder, Matthew. 'Independent Legal Review of the Governance of Biometric Data in England and Wales ('The Ryder Review')', Ada Lovelace Institute (2022), <https://www.adalovelaceinstitute.org/wp-content/uploads/2022/06/The-Ryder-Review-Independent-legal-review-of-the-governance-of-biometric-data-in-England-and-Wales-Ada-Lovelace-Institute-June-2022.pdf>. [中文题名: 英格兰与威尔士生物识别数据治理的独立法律审查 (“Ryder审查”)。]
- Sarat, Austin and Kearns, Thomas (eds) *Law in Everyday Life*, Ann Arbor: University of Michigan Press, 1995. [中文题名: 日常生活中的法律。]
- Scottish Parliament Justice Sub-Committee on Policing. 'Facial Recognition: How Policing in Scotland Makes Use of This Technology', SP Paper 678, 1st Report (Session 5), 2020, <https://sp-bpr-en-prod-cdnep.azureedge.net/published/JSP/2020/2/11/Facial-recognition-how-policing-in-scotland-makes-use-of-this-technology/JSPS0520R01.pdf>. [中文题名: 人脸识别: 苏格兰警务如何使用这项技术。]
- Shahi, Afshin and Abdoh-Tabrizi, Ehsan. 'Iran' s 2019–2020 Demonstrations: The Changing Dynamics of Political Protests in Iran', *Asian Affairs* 51.1 (2020): 1–41. [中文题名: 伊朗2019—2020年示威: 伊朗政治抗议动态的变化。]
- Sharma, Suneet. 'Case Law: R (Bridges) v Chief Constable of South Wales Police: The Use of Facial Recognition Software by the Police is Lawful', *Inform blog*, 6 September 2019, <https://inform.org/2019/09/06/case-law-r-bridges-v-chief-constable-of-south-wales-police-the-use-of-facial-recognition-software-by-the-police-is-lawful-suneet-sharma/>. [中文题名: 判例法: R (Bridges) 诉南威尔士警察总长案——警方使用人脸识别软件合法。]

- U.K. Government. 'National AI Strategy', 21 September 2021, <https://www.gov.uk/government/publications/national-ai-strategy>. [中文题名：国家AI战略。]
- U.K. Government. 'Establishing a Pro-innovation Approach to Regulating AI', 20 July 2022, <https://www.gov.uk/government/publications/establishing-a-pro-innovation-approach-to-regulating-ai>. [中文题名：建立促进创新的AI监管进路。]
- UKRI Economic and Social Research Council. 'Defining Impact', <https://www.ukri.org/councils/esrc/impact-toolkit-for-economic-and-social-sciences/defining-impact/>. [中文题名：定义影响。]
- UNESCO. 'UNESCO Member States Adopt the First Ever Global Agreement on the Ethics of Artificial Intelligence', 25 November 2021, <https://en.unesco.org/news/unesco-member-states-adopt-first-ever-global-agreement-ethics-artificial-intelligence>. [中文题名：联合国教科文组织成员国通过史上首项人工智能伦理全球协议。]
- Urquhart, Lachlan and Miranda, Diana. 'Policing faces: the present and future of intelligent facial surveillance', *Information & Communications Technology Law*, 31:2 (2022): 194-219. [中文题名：面孔警务：智能面部监控的现在与未来。]
- Veale, Michael and Zuiderveen Borgesius, Frederik. 'Demystifying the Draft EU Artificial Intelligence Act—Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach', *Computer Law Review International* 22.4 (2021): 97-112. [中文题名：揭开欧盟人工智能法草案的神秘面纱——分析拟议进路中的良好、不良和不确定要素。]
- Wagner, Ben. 'Ethics as an Escape from Regulation: From 'Ethics-Washing' to Ethics-Shopping?' in Bayamlioglu, Emre, Baraliuc, Irina, Janssens, Liisa, and Hildebrandt, Mireille (eds) *Being Profiled: Cogitas Ergo Sum: 10 Years of Profiling the European Citizen*, Amsterdam: Amsterdam University Press, 2018, pp. 84-9. [中文题名：作为逃避监管手段的伦理：从伦理洗白到伦理选购？；载《被画像：你思故我在——为欧洲公民画像十年》。]
- Weale, Sally. 'ICO to Step in After Schools Use Facial Recognition to Speed Up Lunch Queue', *Guardian*, 18 October 2021, <https://www.theguardian.com/education/2021/oct/18/privacy-fears-as-schools-use-facial-recognition-to-speed-up-lunch-queue-ayrshire-technology-payments-uk>. [中文题名：学校使用人脸识别加快午餐排队后，ICO将介入。]
- Weitzberg, Keren. 'A Very British Problem: The Evolution of Britain's Militarised Policing Industrial Complex', Report for Campaign Against the Arms Trade and Netpol, 2022, <https://caat.org.uk/app/uploads/2022/08/A-Very-British-Problem-WEB.pdf>. [中文题名：一个非常英国的问题：英国军事化警务工业复合体的演变。]
- White, Nadine. 'Met Police Faces Legal Action Over 'Racist' Gangs Matrix Database', *Independent*, 1 February 2022, <https://www.independent.co.uk/news/uk/home-news/met-police-gangs-matrix-database-b2004293.html>. [中文题名：大都会警方因“种族主义”帮派矩阵数据库面临法律行动。]
- Willim, Robert. 'Imperfect Imaginaries: Digitisation, Mundanisation, and the Ungraspable' in Gertraud Koch (ed) *Digitisation: Theories and Concepts for Empirical*

- Cultural Research*, Abingdon: Routledge, 2017, pp. 53–77. [中文题名：不完美的想象：数字化、日常化与不可把握之物；载《数字化：实证文化研究的理论与概念》。]
- Wong, Janis. ‘Protests Decentralised: How Technology Enabled Civil Disobedience by Hong Kong Anti-extradition Bill Protesters’, *LawArXiv*, 2020, <https://osf.io/preprints/lawarxiv/efvwn/>. [中文题名：抗议的去中心化：技术如何促成香港反引渡法案抗议者的公民不服从。]
- Wong, Pak-Hang. ‘Cultural Differences as Excuses? Human Rights and Cultural Values in Global Ethics and Governance of AI’, *Philosophy and Technology* 33 (2020): 705–15. [中文题名：把文化差异当作借口？全球AI伦理与治理中的人权和文化价值。]
- Zalnieriute, Monika. ‘Burning Bridges: The Automated Facial Recognition Technology and Public Space Surveillance in the Modern State’, *Science and Technology Law Review*, 22.2 (2021): 284–307. [中文题名：焚毁桥梁：现代国家中的自动人脸识别技术与公共空间监控。]

原注

- [1] Pak-Hang Wong, ‘Cultural Differences as Excuses? Human Rights and Cultural Values in Global Ethics and Governance of AI’, *Philosophy and Technology* 33 (2020): 705–715. [中文题名：把文化差异当作借口？全球AI伦理与治理中的人权和文化价值。] 返回
- [2] Seán ÓhÉigeartaigh, Jess Whittlestone, Yang Liu, Yi Zeng and Zhe Liu, ‘Overcoming Barriers to Cross-cultural Cooperation in AI Ethics and Governance’, *Philosophy and Technology* 33 (2020): 571–593. [中文题名：克服AI伦理与治理跨文化合作的障碍。] 返回
- [3] 感谢 Xaroula Kerasidou 提出这一点。返回
- [4] UKRI Economic and Social Research Council, ‘Defining Impact’, <https://www.ukri.org/councils/esrc/impact-toolkit-for-economic-and-social-sciences/defining-impact/>. [中文题名：定义影响。] 返回
- [5] Jennifer Chubb and Mark Reed, ‘The Politics of Research Impact: Academic Perceptions of the Implications for Research Funding, Motivation and Quality’, *British Politics* 13 (2018): 302. [中文题名：研究影响的政治：学者对其在研究资助、动机和质量方面影响的看法。] 返回
- [6] Thao Phan, Jake Goldenfein, Monique Mann and Declan Kuch, ‘Economies of Virtue: The Circulation of ‘Ethics’ in Big Tech’, *Science as Culture* 31.1 (2022): 121–35. [中文题名：美德经济：大型科技企业中“伦理”的流通。] 返回
- [7] Mike Finn, *British Universities in the Brexit Moment: Political, Economic and Cultural Implications*, Bingley: Emerald Publishing, 2018, p. 97. [中文题名：脱欧时刻的英国大学：政治、经济与文化含义。] 返回
- [8] 参见本书中 Cath 与 Keyes、Pink，以及 Richardson 的文章。返回
- [9] Rosalind Edwards, ‘Why do Academics do Unfunded Research? Resistance, Compliance and Identity in the UK Neo-liberal University’, *Studies in Higher*

- Education*, 47.4 (2022): 912. [中文题名: 学者为什么开展无资助研究? 英国新自由主义大学中的抵抗、服从与身份。] 返回
- [10] Edwards, ‘Why do academics do unfunded research?’, 912. [中文题名: 学者为什么开展无资助研究? 英国新自由主义大学中的抵抗、服从与身份。] 返回
- [11] U.K. Government, ‘National AI Strategy’, 21 September 2021, <https://www.gov.uk/government/publications/national-ai-strategy>; U.K. Government, ‘Establishing a pro-innovation approach to regulating AI’, 20 July 2022, <https://www.gov.uk/government/publications/establishing-a-pro-innovation-approach-to-regulating-ai>. [中文题名: 国家AI战略; 建立促进创新的AI监管进路。] 返回
- [12] Emre Kazim, Denise Almeida, Nigel Kingsman, Charles Kerrigan, Adriano Koshiyama, Elizabeth Lomas and Airlie Hilliard, ‘Innovation and Opportunity: Review of the UK’s National AI Strategy’, *Discover Artificial Intelligence* 1.14 (2021): 1–10. [中文题名: 创新与机遇: 评述英国国家AI战略。] 返回
- [13] Marinus Ossewaarde and Erdener Gulenc, ‘National Varieties of Artificial Intelligence Discourses: Myth, Utopianism, and Solutionism in West European Policy Expectations’, *Computer* 53.11 (2020): 53–61. [中文题名: 人工智能话语的国家差异: 西欧政策预期中的神话、乌托邦主义与解决主义。] 返回
- [14] Ossewaarde and Gulenc, ‘National Varieties of Artificial Intelligence Discourses’. [中文题名: 人工智能话语的国家差异: 西欧政策预期中的神话、乌托邦主义与解决主义。] 返回
- [15] Thilo Hagendorff, ‘Blind Spots in AI ethics’, *AI Ethics* (2021). [中文题名: AI伦理的盲点。] 返回
- [16] Thilo Hagendorff, ‘The Ethics of AI Ethics: An Evaluation of Guidelines’, *Minds & Machines* 30 (2020), 99–120, 105. [中文题名: AI伦理的伦理: 指南评估。] 返回
- [17] Hagendorff, ‘Blind Spots in AI ethics’. [中文题名: AI伦理的盲点。] 返回
- [18] 感谢 Jake Goldenfein 提出这一点。返回
- [19] 感谢 Jake Goldenfein 提出这一点。返回
- [20] 感谢 Jake Goldenfein 提出这一点。返回
- [21] Jean Burgess, Peta Mitchell and Tim Highfield, ‘Automating the Digital Everyday: An Introduction’, *Media International Australia*, 166.1 (2018): 6–10, 6. [中文题名: 日常数字生活的自动化: 导论。] 返回
- [22] Pink, Ruckenstein, Berg and Lupton, ‘Everyday Automation: Setting a Research Agenda’. [中文题名: 日常自动化: 设定研究议程; 载《日常自动化: 体验并预期新兴技术》。] 返回
- [23] 例如参见, Anna Jobin, Marcello Ienca and Effy Vayena, ‘The Global Landscape of AI Ethics Guidelines’, *Nature Machine Intelligence* 1 (2019): 389–99. [中文题名: AI伦理指南的全球格局。] 返回

- [24] Pink, Ruckenstein, Berg and Lupton, 'Everyday Automation: Setting a Research Agenda', 8. [中文题名: 日常自动化: 设定研究议程; 载《日常自动化: 体验并预期新兴技术》。] 返回
- [25] Alex Jiahong Lu, 'Toward Everyday Negotiation and Resistance Under Data-Driven Surveillance', *Interactions* 29.2 (2022). [中文题名: 迈向数据驱动监控下的日常协商与抵抗。] 返回
- [26] Ben Wagner, 'Ethics as an Escape from Regulation: From 'Ethics-Washing' to Ethics-Shopping?', in Emre Bayamlioglu, Irina Baraliuc, Liisa Janssens and Mireille Hildebrandt (eds) *Being Profiled: Cogitas Ergo Sum: 10 Years of Profiling the European Citizen*, Amsterdam: Amsterdam University Press, 2018, pp. 84–9. [中文题名: 作为逃避监管手段的伦理: 从伦理洗白到伦理选购? ; 载《被画像: 你思故我在——为欧洲公民画像十年》。] 返回
- [27] Elettra Bietti, 'From Ethics Washing to Ethics Bashing: A View on Tech Ethics from Within Moral Philosophy', *Proceedings of ACM FAT* Conference*, 2020, <https://ssrn.com/abstract=3513182>. [中文题名: 从伦理洗白到伦理抨击: 道德哲学内部的科技伦理观。] 返回
- [28] Angela Daly, S. Kate Devitt and Monique Mann, 'AI Ethics Needs Good Data', in Pieter Verdegem (ed) *AI for Everyone? Critical Perspectives*, London: University of Westminster Press, 2021. [中文题名: AI伦理需要良好数据; 载《人人共享AI? 批判性视角》。] 返回
- [29] Angela Daly, Thilo Hagendorff, Li Hui, Monique Mann, Vidushi Marda, Ben Wagner and Wayne Wei Wang, 'AI, Governance and Ethics: Global Perspectives' in Hans Micklitz, Oreste Pollicino, Amnon Reichman, Andrea Simoncini, Giovanni Sartor and Giovanni De Gregorio (eds) *Constitutional Challenges in the Algorithmic Society*, Cambridge: Cambridge University Press, 2022. [中文题名: AI、治理与伦理: 全球视角; 载《算法社会中的宪制挑战》。] 返回
- [30] UNESCO, 'UNESCO member states adopt the first ever global agreement on the Ethics of Artificial Intelligence', 25 November 2021, <https://en.unesco.org/news/unesco-member-states-adopt-first-ever-global-agreement-ethics-artificial-intelligence>. [中文题名: 联合国教科文组织成员国通过史上首项人工智能伦理全球协议。] 返回
- [31] 例如参见, Michael Veale and Frederik Zuiderveen Borgesius, 'Demystifying the Draft EU Artificial Intelligence Act—Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach', *Computer Law Review International* 22.4 (2021): 97–112. [中文题名: 揭开欧盟人工智能法案的神秘面纱——分析拟议进路中的良好、不良和不明确要素。] 返回
- [32] Leigh McGowran, 'The Issues with the EU's Draft Regulation on Facial Recognition AI', *Silicon Republic*, 17 May 2022, <https://www.siliconrepublic.com/enterprise/the-issues-with-the-eus-draft-regulation-on-facial-recognition-ai>. [中文题名: 欧盟关于人脸识别AI的法规草案存在的问题。] 返回

- [33] 例如参见, Patricia Ewick and Susan Silbey, *The Common Place of Law: Stories from Everyday Life*, Chicago: University of Chicago Press, 1998; Richard Moule, George Burruss, Faith Gifford, Megan Parry and Bryanna Fox, 'Legal Socialization and Subcultural Norms: Examining Linkages Between Perceptions of Procedural Justice, Legal Cynicism, and the Code of the Street', *Journal of Criminal Justice* 61 (2019): 26–39. [中文题名: 法律的日常场所: 日常生活故事; 法律社会化与亚文化规范: 考察程序正义感知、法律犬儒主义和街头守则之间的联系。] 返回
- [34] Austin Sarat and Thomas Kearns (eds) *Law in Everyday Life*. Ann Arbor: University of Michigan Press, 1995, p. 9. [中文题名: 日常生活中的法律。] 返回
- [35] Michael McCann, 'Law and Social Movements: Contemporary Perspectives', *Annual Review of Law and Social Science* 2.1 (2006): 17–38. [中文题名: 法律与社会运动: 当代视角。] 返回
- [36] Becky Kazansky, Guillen Torres, Lonneke van der Velden, Kersti Wissenbach, and Stefania Milan, 'Data for the Social Good: Towards a Data-Activist Research Agenda', in Angela Daly, S. Kate Devitt and Monique Mann (eds), *Good Data*, Amsterdam: Institute of Network Cultures, 2019, 246. [中文题名: 服务社会公益的数据: 迈向数据行动主义研究议程; 载《良好数据》。] 返回
- [37] Davide Beraldo and Stefania Milan, 'From Data Politics to the Contentious Politics of Data', *Big Data & Society* 6.2 (2019): 2. [中文题名: 从数据政治到数据的争议政治。] 返回
- [38] Becky Kazansky, '“It Depends on your Threat Model”: The Anticipatory Dimensions of Resistance to Data-driven Surveillance', *Big Data & Society* 8.1 (2021): 1. 另见 Lu, 'Toward Everyday Negotiation and Resistance Under Data-Driven Surveillance'. [中文题名: “取决于你的威胁模型”: 抵抗数据驱动监控的预期性维度; 迈向数据驱动监控下的日常协商与抵抗。] 返回
- [39] Anne Kaun, 'Suing the Algorithm: The Mundanization of Automated Decision-making in Public Services Through Litigation', *Information, Communication & Society* (2021); Robert Willim, 'Imperfect imaginaries: Digitisation, mundanisation, and the ungraspable' in Gertraud Koch (ed), *Digitisation: Theories and Concepts for Empirical Cultural Research*, Abingdon: Routledge, 2017. [中文题名: 起诉算法: 通过诉讼使公共服务中的自动化决策日常化; 不完美的想象: 数字化、日常化与不可把握之物; 载《数字化: 实证文化研究的理论与概念》。] 返回
- [40] Monique Mann, 'Social (In)security and Social (In)justice: Automation in the Australian Welfare System' in *Artificial Intelligence: Human Rights, Social Justice and Development: Global Information Society Watch 2019 Report*, 2019, pp. 68–72. [中文题名: 社会(不)安全与社会(不)正义: 澳大利亚福利系统中的自动化; 载《人工智能: 人权、社会正义与发展——2019年全球信息社会观察报告》。] 返回
- [41] Tulsi Parida and Aparna Ashok, 'Consolidating Power in the Name of Progress: Techno-solutionism and Farmer Protests in India' in Frederike Kaltheuner (ed), *Fake AI*, Manchester: Meatspace Press, 2021, pp. 161–9. [中文题名: 以进步之名巩固权力: 印度的技术解决主义与农民抗议; 载《虚假AI》。] 返回

- [42] Joy Buolamwini and Timnit Gebru, 'Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification' in *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, PMLR 81, 2018, 1. [中文题名: 性别色差: 商业性别分类中交叉性的准确率差异。] 返回
- [43] Angela Daly, 'Algorithmic Oppression with Chinese Characteristics: AI Against Xinjiang's Uyghurs' in *Artificial Intelligence: Human Rights, Social Justice and Development: Global Information Society Watch 2019 Report*, 2019, pp. 108–12. [中文题名: 具有中国特色的算法压迫: 针对新疆维吾尔人的AI; 载《人工智能: 人权、社会正义与发展——2019年全球信息社会观察报告》。] 返回
- [44] Isadora Neroni Rezende, 'Facial Recognition in Police Hands: Assessing the 'Clearview case' from a European Perspective', *New Journal of European Criminal Law* 11.3 (2020): 375–89. [中文题名: 警方手中的人脸识别: 从欧洲视角评估“Clearview案”。] 返回
- [45] Melissa Heikkilä, 'The Walls are Closing in on Clearview AI', *MIT Technology Review*, 24 May 2022, <https://www.technologyreview.com/2022/05/24/1052653/clearview-ai-data-privacy-uk/>. [中文题名: Clearview AI四面楚歌。] 返回
- [46] Afshin Shahi and Ehsan Abdoh-Tabrizi, 'Iran' s 2019–2020 Demonstrations: The Changing Dynamics of Political Protests in Iran', *Asian Affairs* 51.1 (2020): 1–41. [中文题名: 伊朗2019—2020年示威: 伊朗政治抗议动态的变化。] 返回
- [47] 例如参见, @DrParchizadeh, 'Protesters in Tehran sabotage the police CCTV so that they can' t be identified, arrested and killed by the regime. #IranProtests', Twitter post, 16 November 2019, 4:25PM, <https://twitter.com/DrParchizadeh/status/1195739605460496385>. [中文题名: 德黑兰抗议者破坏警方闭路电视摄像头, 使政权无法识别、逮捕并杀死他们。#IranProtests [推文, 2019年11月16日下午4:25。]] 返回
- [48] @javidirani30, 'Last night, Monday, July 19th, Ahwazi youths in Alavi alley disabled CCTV cameras #Khuzestan #IranProtests', Twitter post, 20 July 2021, 8:45AM, <https://twitter.com/javidirani30/status/1417390100720279569>. [中文题名: 昨夜, 7月19日星期一, 阿瓦兹青年在Alavi巷使闭路电视摄像头失效。#Khuzestan #IranProtests [推文, 2021年7月20日上午8:45。]] 返回
- [49] Azadeh Akbari, 'The Threat of Automating Control: Surveillance of Women' s Clothing in Iran', in Aleš Završnik and Vasja Badalič (eds) *Automating Crime Prevention, Surveillance, and Military Operations*, Cham: Springer, 2021, 186. [中文题名: 自动化控制的威胁: 伊朗对女性服装的监控; 载《犯罪预防、监控与军事行动的自动化》。] 返回
- [50] Janis Wong, 'Protests Decentralised: How Technology Enabled Civil Disobedience by Hong Kong Anti-extradition Bill Protesters', *LawArXiv*, 2020, <https://osf.io/preprints/lawarxiv/efvwn/>. [中文题名: 抗议的去中心化: 技术如何促成香港反引渡法案抗议者的公民不服从。] 返回
- [51] Manoj Kewalramani and Rohan Seth, 'Networked Protests & State Responses: The Case of Hong Kong 2019–2020', *Takshashila Discussion Document 2020-03*, 2020,

- <https://ssrn.com/abstract=3580591>. [中文题名: 网络化抗议与国家回应: 2019—2020 年的香港案例。] 返回
- [52] Wong, ‘Protests Decentralised’, 6. [中文题名: 抗议的去中心化: 技术如何促成香港反引渡法案抗议者的公民不服从。] 返回
- [53] ‘Hong Kong: Anti-surveillance Protestors Tear Down ‘Smart’ Lamp-post’, *Guardian*, 26 August 2019, <https://www.theguardian.com/world/video/2019/aug/26/hong-kong-anti-surveillance-protesters-tear-down-smart-lamp-post-video>. [中文题名: 香港: 反监控抗议者拆倒“智能”灯柱。] 返回
- [54] Sean Gleeson, ‘How Smart are Hong Kong’s Lampposts?’, *AFP Fact Check*, 4 September 2019, <https://factcheck.afp.com/how-smart-are-hong-kongs-lampposts>. [中文题名: 香港的灯柱有多智能?。] 返回
- [55] Kevin Leung and H.Y. Lee, ‘Implementing the Smart City: Who Has a Say? Some Insights from Hong Kong’, *International Journal of Urban Sciences*, 2021, 1–25. [中文题名: 实施智慧城市: 谁有发言权? 来自香港的一些观察。] 返回
- [56] Dave Lee, ‘San Francisco is First US City to Ban Facial Recognition’, *BBC*, 15 May 2019, <https://www.bbc.co.uk/news/technology-48276660>. [中文题名: 旧金山成为美国首个禁止人脸识别的城市。] 返回
- [57] Todd Feathers, ‘Facial Recognition Is Racist. Why Aren’t More Cities Banning It?’, *Vice*, 25 May 2021, <https://www.vice.com/en/article/4avx3m/facial-recognition-is-racist-why-arent-more-cities-banning-it>. [中文题名: 人脸识别具有种族歧视性, 为什么没有更多城市禁止它?。] 返回
- [58] Niall Hamilton-Smith, Maureen McBride and Colin Atkinson, ‘Lights, Camera, Provocation? Exploring Experiences of Surveillance in the Policing of Scottish Football’, *Policing and Society* 31.2 (2021): 179–94. [中文题名: 灯光、摄像、挑衅? 探究苏格兰足球警务中的监控体验。] 返回
- [59] Graham Ruthven, ‘The Criminalization of Scottish Soccer Fans’, *Vice*, 23 February 2016, <https://www.vice.com/en/article/9apyad/the-criminalization-of-scottish-soccer-fans>. [中文题名: 苏格兰球迷的犯罪化。] 返回
- [60] Marcello Mega, ‘Cops Fear Gangsters are Evading Law as Glasgow’s Facial Recognition Cameras Remain Mothballed’, *Daily Record*, 11 August 2020, <https://www.dailyrecord.co.uk/news/scottish-news/scots-cops-fear-gangsters-evading-22499468>. [中文题名: 格拉斯哥人脸识别摄像头仍被封存, 警察担忧黑帮分子逃避法律。] 返回
- [61] Keren Weitzberg, ‘A Very British Problem: The Evolution of Britain’s Militarised Policing Industrial Complex’, Report for Campaign Against the Arms Trade and Netpol, 2022, <https://caat.org.uk/app/uploads/2022/08/A-Very-British-Problem-WEB.pdf>. [中文题名: 一个非常英国的问题: 英国军事化警务工业复合体的演变。] 返回
- [62] Scottish Parliament Justice Sub-Committee on Policing, ‘Facial Recognition: How Policing in Scotland Makes Use of This Technology’, SP Paper 678, 1st Report (Session 5), 2020, <https://sp-bpr-en-prod-cdnep.azureedge.net/published/JSP/2020/2>

- /11/Facial-recognition--how-policing-in-Scotland-makes-use-of-this-technology/JSPS0520R01.pdf. [中文题名: 人脸识别: 苏格兰警务如何使用这项技术。] 返回
- [63] Big Brother Watch, ‘Face Off: The Lawless Growth of Facial Recognition in UK Policing’, 2018, <http://bigbrotherwatch.org.uk/wp-content/uploads/2018/05/Face-Off-final-digital-1.pdf>. [中文题名: 面孔对抗: 英国警务中人脸识别不受法律约束的扩张。] 返回
- [64] 例如参见, Lisa Long and Remi Joseph-Salisbury, ‘Black Mixed-race Men’s Perceptions and Experiences of the Police’, *Ethnic and Racial Studies* 42.2 (2019): 198–215. [中文题名: 黑人混血男性对警方的看法与经历。] 返回
- [65] Jasbinder Nijjar, ‘Police–school Partnerships and the War on Black Youth’, *Critical Social Policy* 41.3 (2021): 491–501. [中文题名: 警校合作与针对黑人青年的战争。] 返回
- [66] Nadine White, ‘Met Police Faces Legal Action Over ‘Racist’ Gangs Matrix Database’, *Independent*, 1 February 2022, <https://www.independent.co.uk/news/uk/home-news/met-police-gangs-matrix-database-b2004293.html>. [中文题名: 大都会警方因“种族主义”帮派矩阵数据库面临法律行动。] 返回
- [67] Lizzie Dearden, ‘Police Stop People for Covering their Faces from Facial Recognition Camera Then Fine Man £90 After he Protested’, *The Independent*, 31 January 2019, <https://www.independent.co.uk/news/uk/crime/facial-recognition-cameras-technology-london-trial-met-police-face-cover-man-fined-a8756936.html>. [中文题名: 警方拦下为躲避人脸识别摄像头而遮脸的人, 一名男子抗议后被罚90英镑。] 返回
- [68] Weitzberg, ‘A Very British Problem’. [中文题名: 一个非常英国的问题: 英国军事化警务工业复合体的演变。] 返回
- [69] Ada Lovelace Institute, ‘Beyond Face Value: Public Attitudes to Facial Recognition Technology’, 2019, <https://www.adalovelaceinstitute.org/report/beyond-face-value-public-attitudes-to-facial-recognition-technology/>. [中文题名: 不止于表面: 公众对人脸识别技术的态度。] 返回
- [70] Lachlan Urquhart and Diana Miranda, ‘Policing Faces: The Present and Future of Intelligent Facial Surveillance’, *Information & Communications Technology Law* 31.2 (2022): 194–219, 198. [中文题名: 面孔警务: 智能面部监控的现在与未来。] 返回
- [71] Monika Zalnieriute, ‘Burning Bridges: The Automated Facial Recognition Technology and Public Space Surveillance in the Modern State’, *Science and Technology Law Review*, 22.2 (2021): 284–307, 287. [中文题名: 焚毁桥梁: 现代国家中的自动人脸识别技术与公共空间监控。] 返回
- [72] Suneet Sharma, ‘Case Law: R (Bridges) v Chief Constable of South Wales Police: The Use of Facial Recognition Software by the Police is Lawful’, *Inform blog*, 6 September 2019, <https://inform.org/2019/09/06/case-law-r-bridges-v-chief-constable-of-south-wales-police-the-use-of-facial-recognition-software-by-the-police-is-lawful-suneet-sharma/>. [中文题名: 判例法: R (Bridges) 诉南威尔士警察总长案——警方使用人脸识别软件合法。] 返回

- [73] Hunton Andrews Kurth, ‘UK Court of Appeal Finds Automated Facial Recognition Technology Unlawful in Bridges v South Wales Police’ 12 August 2020, <https://www.huntonprivacyblog.com/2020/08/12/uk-court-of-appeal-finds-automated-facial-recognition-technology-unlawful-in-bridges-v-south-wales-police/>; 另见 Urquhart and Miranda, ‘Policing Faces’ . [中文题名: 英国上诉法院在Bridges诉南威尔士警方案中认定自动人脸识别技术违法; 面孔警务: 智能面部监控的现在与未来。] 返回
- [74] Zalnieriute, ‘Burning Bridges’ . [中文题名: 焚毁桥梁: 现代国家中的自动人脸识别技术与公共空间监控。] 返回
- [75] Euan Lynch, ‘The Use of Live Facial Recognition Technology in Scotland: A New North–South Divide?’ , *UK Human Rights blog*, 25 February 2020, <https://ukhumanrightsblog.com/2020/02/25/the-use-of-live-facial-recognition-technology-in-scotland-a-new-north-south-divide/>. [中文题名: 苏格兰实时人脸识别技术的使用: 新的南北分野? 。] 返回
- [76] Sally Weale, ‘ICO to Step In After Schools Use Facial Recognition to Speed Up Lunch Queue’ , *Guardian*, 18 October 2021, <https://www.theguardian.com/education/2021/oct/18/privacy-fears-as-schools-use-facial-recognition-to-speed-up-lunch-queue-ayrshire-technology-payments-uk>. [中文题名: 学校使用人脸识别加快午餐排队后, ICO将介入。] 返回
- [77] Pascale Davies, ‘UK Schools Suspend Use of Controversial Facial Recognition Technology’ , *Euronews*, 25 October 2021, <https://www.euronews.com/next/2021/10/18/schools-in-scotland-start-using-facial-recognition-on-children-paying-for-lunch>. [中文题名: 英国学校暂停使用引发争议的人脸识别技术。] 返回
- [78] Asssiah Hamed, ‘Co-op Defends Facial Recognition Cameras in Bristol Stores Amid Claims of ‘Orwellian’ Ssurveillance’ , *Bristol Post*, 9 December 2021, <https://www.bristolpost.co.uk/whats-on/shopping/co-op-defends-facial-recognition-6302476>. [中文题名: 被指实施“奥威尔式”监控之际, Co-op为布里斯托门店的人脸识别摄像头辩护。] 返回
- [79] ‘Southern Co-operative’ s Use of Facial Recognition on Customers Prompts Legal Complaint’ , *Sky News*, 27 July 2022, <https://news.sky.com/story/co-ops-use-of-facial-recognition-on-customers-prompts-legal-complaint-12659309>. [中文题名: Southern Co-operative对顾客使用人脸识别引发法律投诉。] 返回
- [80] Matthew Ryder, ‘Independent Legal Review of the Governance of Biometric Data in England and Wales (‘The Ryder Review’)’ , Ada Lovelace Institute, 2022, <https://www.adalovelaceinstitute.org/wp-content/uploads/2022/06/The-Ryder-Review-Independent-legal-review-of-the-governance-of-biometric-data-in-England-and-Wales-Ada-Lovelace-Institute-June-2022.pdf>. [中文题名: 英格兰与威尔士生物识别数据治理的独立法律审查 (“Ryder审查”) 。] 返回
- [81] Anu Bradford, *The Brussels Effect: How the European Union Rules the World*. Oxford: Oxford University Press, 2020. [中文题名: 布鲁塞尔效应: 欧盟如何统治世界。] 返回
- [82] Angela Daly, ‘Neo-Liberal Business-As-Usual or Post-Surveillance Capitalism With European Characteristics? The EU’ s General Data Protection Regulation in a Multi-

Polar Internet’ , in Rolien Hoyng and Gladys Pak Lei Chong (eds) *Critiquing Communication Innovation: New Media in a Multipolar World*, East Lansing: Michigan State University Press, 2022 (forthcoming). [中文题名: 新自由主义照常运转, 还是具有欧洲特色的后监控资本主义? 多极互联网中的欧盟《通用数据保护条例》; 载《批判传播创新: 多极世界中的新媒体》(即将出版)。] [返回](#)

[83] Ossewaarde and Gulenc, ‘National Varieties of Artificial Intelligence Discourses’ . [中文题名: 人工智能话语的国家差异: 西欧政策预期中的神话、乌托邦主义与解决主义。] [返回](#)

[84] Ian Elliott, Karin Bottom, Paul Carmichael, Joyce Liddle, Steve Martin, and Robert Pyper, ‘The Fragmentation of Public Administration: Differentiated and Decentered Governance in the (dis)United Kingdom’ , *Public Administration* 100.1 (2022): 98–115. [中文题名: 公共行政的碎片化: (不)联合王国中的差异化与去中心化治理。] [返回](#)

以数据为宴： 伦理、价值与风险偏好的计算

Tsvetelina Hristova 和 Liam Magee

数据伦理与AI伦理构成了一个争议日益激烈的领域，学者、行动者、国家机构与产业参与者竞相界定伦理实践的原则。国家认可的伦理框架、由行动者和学者牵头的FAIR数据原则等倡议^[1]，以及微软Aether委员会等产业项目，虽已得到广泛讨论，但国际标准作为影响数据伦理如何被理解和实践的治理技术之一，仍大体处于研究者的关注范围之外。本章考察一系列相互关联的风险管理标准的作用；这些标准在塑造一种特定的数据伦理观方面，已经发挥越来越重要的作用。

我们先介绍其中几项全球标准；它们彼此之间以及与先行国家标准之间的联系，可能难以厘清。这些标准以澳大利亚和新西兰风险管理标准AS/NZS 4360-2004为基础；后者也是国际风险管理标准ISO 31000的基础。它们将风险与价值之间的一种特定关系编纂为规范。这种关系凸显出国家和产业参与者如何想象数据使用的社会维度，以及自身在全球数字经济中的角色。澳大利亚标准AS/NZS 4360-2004与ISO 31000引入了一套新的风险管理框架，将风险概念化为具有两面性：在某些情况下为正面，在另一些情况下则为负面。风险的两面性催生了一种以组织和机构的“风险偏好”为依据的风险管理方法——即根据预期收益而形成的承担风险的意愿。这套风险管理框架被纳入新设的ISO分技术委员会ISO/IEC JTC 1/SC 42的工作中，该委员会负责制定人工智能标准。ISO/IEC JTC 1/SC 42属于一个规模更大的AI标准制定倡议；该倡议由信息与通信技术标准化联合技术委员会JTC 1的一个专门分技术委员会牵头。JTC 1成立于1987年，旨在整合两大国际标准机构——国际标准化组织（ISO）和国际电工委员会（IEC）——的标准化工作。ISO/IEC JTC 1/SC 42分技术委员会则于较近的2018年成立，受命制定人工智能领域的标准。这一项目的重要性不容低估：IEC，尤其是ISO，已经建立起学者所谓的“通过共识实现的全球治理”^[2]；它们各自的全球标准经由各种合规要求，影响着大多数行业。在该分技术委员会制定的一系列标准中，小组还在制定AI风险管理标准ISO/IEC DIS 23894；正如其引言所述，该标准明确以ISO 31000为基础：

本文件旨在与ISO 31000:2018配套使用。当本文件扩展ISO 31000:2018所给出的指导时，会适当引用ISO 31000:2018的相应条款；如适用，随后给出AI专门指导。为了使本文件与ISO 31000:2018的关系更加明确，本文件沿用ISO 31000:2018的条款结构，并在必要时增设分条款。^[3]

ISO 31000的案例所提示的不只是标准的可复制性，还包括一种正在形成的社会技术构型：风险开始参与塑造人们如何针对数据主体想象并实施伦理与治理。用户与公司都被赋予一种内在的“风险偏好”；它使大数据与AI情境中不同程度的偶然性变得可以容许，并划定预期和获准的数据实践的边界。数据主体也被卷入技术标准与民族国家地缘政治抱负之间复杂的相互作用之中；在这种相互作用中，地方层面的数据与AI监管旨在成为全球治理框架的试验场。

就这一政治经济空间而言，标准处于一个复杂的位置。Keller Easterling^[4]专门以ISO为例提出“国家之外的治理术”（extrastatecraft）概念，用来刻画重新安排权力、控制和生产关系的技术与经济机制。这些机制并非仅由民族国家政府引导，而且可以重塑国家、城市或大陆内部及彼此之间的政治结构和权力的空间维度。Andrew Barry^[5]提出了类似论点，主张“技术区”这一概念；技术区由复杂的技术基础设施与关系界定，而非由国家政府的传统政治权力界定。在两位学者看来，国家之外的治理术和技术区网络内部各种力量的互动，都由技术基础设施、协议与经济和政治权力之间的相互关联所塑造。标准通过在公司和国家机构之间缔结联盟，并将形成共识的原则和技术约束作为行使控制的方式，占据了这一处在政治治理惯常规范之外的国家之外治理术的规制空间；它们由此塑造出一个容许某些经济流动与关系、同时限制另一些流动与关系的空间。

技术标准与伦理的明确纠缠虽然相对较新，但标准长期以来一直是社会技术组装体^[6]的关键组成部分；数据在这些组装体中得到定义，并被用于不同的计算和统计操作。在历史上，标准在数据基础设施与网络如何被治理^[7]、不同数字文件格式如何被界定^[8]，以及使数据能够跨不同系统流通的互操作性这一关键原则如何被理解与落实^[9]等方面，都发挥了重要作用。标准对于创造大数据收集、交换和分析的条件至关重要。而且，正如我们已经指出的，除了界定数据格式与网络基础设施的作用，标准在更一般的意义上还处于一个模糊空间：它们通过实质上属于“国家之外的治理术”^[10]的共识建构和技术官僚方法，塑造政治与经济进程。我们将ISO 31000这样的标准视为数据与算法的复杂、变动的社会技术组装体的一部分；数据与风险的关系正是在这些组装体中确定的。数据作为风险管理工具体系的一部分，已在用于先发控制与警务的算法研究中得到广泛探讨^[11]；而ISO 31000则通过新的AI标准ISO/IEC DIS 23894，将数据与AI本身置于风险评估和缓解的对象位置。

这些标准对于阐明风险与价值之间的关系至关重要；作为标准建立的社会技术控制体制的一部分，这种关系对于数据使用和伦理如何被想象、机构与国家如何看待自身在大数据与AI治理中的作用，正变得越来越重要。如后文所述，控制从技术延伸至想象层面的一个有趣面向，是“风险”词汇的标准化。例如，在这一过程中，建构出来的“风险偏好”概念及其相关术语被赋予了重要作用；它们共同塑造了一种不同的制度想象，一种已经赋予数据独特的道德能动性 & 经济能动性的想象。

风险以新的方式进入如何治理数字经济并从中获益的盘算，这对不同领域和不同地缘政治行动者如何理解并操作化风险产生了后果。在标准与指导原则等“通过共识进行治理”的技术中，这一点尤其显著；这些技术为实践在私营企业、国家议程以及国际合作与影响力等领域之间的协商和转译留下了空间。我们在此聚焦的风险管理标准ISO 31000，在这一意义上是一个突出的案例：这不仅因为它重申了承担风险与风险偏好的语言对于成功治理不可或缺，也因为这一标准的生命历程揭示了AI伦理与风险指导原则的制定所依托的利害关系和期望。

作为国际标准，ISO 31000成为一种新方式的标志：不同的国家或地方行动者通过制定标准与伦理规范并为之游说，在数据密集型产业中开辟影响力和领导地位。它表明，风险及其不断变化的意义与功能，不仅通过风险偏好原则的采纳发挥变革作用，还通过处于标准制定倡议核心的特定地缘政治和地缘经济抱负发挥作用。与把技术标准理解为以测量、科学与理性的客观性为基础的日常看法相反，标准确实在塑造并表达地缘政治和地缘经济的抱负与边界方面发挥着重要作用。即使标准由非政府机构制定，其地理来源及背后组织的构成，也被视为代表特定的国家利益。正如我们在别处研究技术标准被用于地缘政治和地缘经济斗争的案例时所说明的^[12]，国家政治权力与技术区经常相互纠缠，共同塑造并重塑一项标准的影响范围。认为某些国家能够通过其所支持的标准获得国际采纳而扩大影响力，这种看法本身正是这些纠缠付诸实践的重要组成部分。

新数字经济中的风险标准化

传统上，风险根据其应用的尺度与场所，被概念化为多种不同形式。一种得到广泛接受的定义来自灾害管理：风险是危险性、暴露度与脆弱性的组合。^[13]组织风险常以类似的方式定义：它是可能性（可比于危险性）与严重程度（结合了暴露度和脆弱性）的组合。但风险被认为在社会政治生活中发挥着深刻得多的作用。例如，Georg Simmel在《货币哲学》中，将风险与抽象经济价值及社会层面的信任之间的一种根本关系联系起来。他认为^[14]，在货币和信用经济中，由于不确定性（或社会风险）这一潜在因素的存在，信任通过构建一种社交伦理的必要情境，成为价值生产不可分割的一部分；这种情境能够容纳并抵消冒险的危险。风险后来成为组织明确关注的对象，Ulrich Beck^[15]的论述反映了这一发展：他认为，风险管理已发展为后现代时代治理的决定性特征。他指出，风险地位取代阶级地位，成为当代生存的关键；风险的生产、管理和控制，也已经取代政府早先对价值及其分配的关切。

过去二十年间，在管理话语及与ISO 31000有关的标准家族中，风险如今表现为具有两面性、而非纯粹负面的存在。ISO 31000建立了一套术语，将风险定义为“不确定性对目标的影响”，并进一步解释为“偏离预期的情况 [……这种偏离] 可以是正面的、负面的或两者兼具，并可以应对、创造或导致机会与威胁”。^[16]得益

于国际标准化组织的重大影响力，这种解释得到广泛采纳；它把风险描绘为一个能够为公司和机构带来收益与价值的不确定性领域。风险的诱惑力再次肯定了该标准 为风险管理原则所设定的核心目标：创造和保护价值。

如果风险创造机会，就有必要追求冒险。配套的ISO Guide 73:2009通过“风险偏好”概念表述了这一要求，即“组织愿意追求或保留的风险数量和类型”。^[17]在全球咨询公司德勤^[18]随后对风险偏好的阐释中，消费的隐喻被进一步推展，并融入一个区分不同风险程度的框架：公司承担的风险数量，会使其处于饥饿、饱足或过饱状态。这种消化隐喻无论多么牵强，仍提醒我们注意历史上将经济组织构想为身体的做法：公司也是一个有形的身体，有生命、会呼吸，需要自身的法律保护，并在生产的同时消费。即使大学风险管理声明的语言，也能说明其治理与赌博消费的决策框架有多么相似：先计算胜算，再根据组织的偏好，押注于低风险或高风险结果。

全球标准ISO 31000采纳风险偏好和风险容忍度这些概念后，它们便成为其他所有风险管理相关标准中可复制的模型，包括信息技术安全管理标准ISO/IEC 27001，以及上述AI风险管理标准。该标准引入了对如何处理风险的新理解，显著改变了国家、企业和国际治理语境中的风险概念及风险管理参数。可接受风险和风险偏好的概念，也开始影响公共机构如何理解自己对国家收集的数据所负的职责。例如，澳大利亚皇家生产力委员会2017年发布的《数据可得性与使用》报告，大力鼓励公私实体之间的数据共享，尤其鼓励共享公共数据集，以促进本地数字产业的经济增长与创新。值得注意的是，它以风险偏好以及社会对共享个人数据相关风险的容忍度提高为依据，论证应当采取更宽松的数据共享方式。报告主张，既然社会的隐私标准已经改变——社交媒体公司使新的数据共享实践常态化，在很大程度上推动了这种转变——联邦机构也应当在共享数据时承担更大的组织风险。

与ISO 31000有关的AI风险管理标准的出现，意味着对数据及其价值的关注更加密集。然而，数据的经济化——我们在此指价值的生产，无论以直接财务收益，还是以间接的声誉、人力资源或政治利益衡量——绝非新鲜事。我们将此处的讨论置于AI在2010年代走出此前数十年所谓“寒冬”的背景之中。据Facebook首席科学家、AI先驱LeCun^[19]所述，这一复兴来自以下要素的结合：硬件（尤其是将GPU改用于并行数据处理）、软件（如Facebook的PyTorch和Google的TensorFlow等开源库）、研究方法的改进（尤其是神经网络、随机梯度下降和反向传播的使用），以及大型文本、图像和其他媒体数据集的积累。对于Facebook、Google、百度和阿里巴巴等科技公司——它们既是AI研究的领军者，也遥遥领先于其他企业，最善于把在线广告转化为收入——数据与价值之间的联系，不仅体现在向消费者投放相关广告的效率上，也体现在它们向消费者提供的一系列AI驱动的服务上，从搜索结果到内容审核皆是如此。在价值方面较不显眼的一点，是这些公司能够吸引LeCun和Geoffrey Hinton（Google）等知名数据科学家，其承诺是让他们有机

会使用海量数据集，解决社会和计算问题，产出“最先进”的AI研究。在算法范式似乎正从庞大的代码库转向以充足数据进行训练的较小代码模型的时期，数据的大量存在不仅创造规模经济，使组织在广告商和消费者面前脱颖而出，而且也是吸引人才的关键因素。

对于资本主义的新巨头，以及其他私营和公共机构，数据治理都是核心业务。数据的存在本身，也给它所创造的价值带来风险。数据可能遭到窃取、泄露、滥用、损坏、忽视和歪曲。虽然数据远非风险管理的唯一对象，但由于它在组织运作中的中心地位，二十世纪和二十一世纪通过程序、审查与标准使风险逐步形式化的过程中，数据同样至关重要。^[20]从著名的安全泄露事件，到利用社交网络关系图谱定向投放政治信息^[21]，各式各样的数据危机也无意间造就了一种公开宣称自身合乎伦理的组织主体；它通过条款、条件、原则、章程、承诺、政策与标准，保证保护其控制范围内的数据。正如我们在此所论证的，这种主体一边将自身伦理作为众多公开展示的美德之一加以陈列，一边也准备让伦理在价值生产中承受风险。因此，制定AI伦理框架的倡议日益增多，并不一定意味着它们在试图将数据伤害的风险降至最低。相反，它们同样可能说明，人们正在转向拥抱这些风险，并使之由社会共同承担。

风险在数据经济与AI监管中被操作化的格局发生变化，意味着我们面对的是另一种确立和实施伦理观念的关系构型。在这个新情境中，数据、自动化与人工智能并不只是强化一种排除疑虑和不确定性的治理方式。^[22]它们也被部署于风险经济之中：风险在一定程度内是可接受的，甚至受到主动追求，因为它打开了不确定性与模糊性的缝隙，并由此带来实现经济收益和其他形式价值的可能。这种刻画呼应了早期关于风险与企业家精神、创新和资本之间关系的理论；例如Brouwer^[23]就讨论了Weber、Schumpeter和Knight对于承担风险的企业家所作的不同刻画。不过，尽管这些理论家各有差异，他们都设想资本主义早期的风险最终会让位于理性主义、垄断性的晚期资本主义组织；而在当下的数据经济中，我们可以看到，恰恰是这些组织自身有意将风险重新注入，作为一种激活机制，以提振日趋疲弱的利润率。换言之，风险既是颠覆性企业家的特征，也是已站稳脚跟、高度理性主义的市场巨头的特征。

随着经济行动者采纳风险偏好，对于数据这样一种高度社会化的资源，伦理成为协商榨取与营利边界的关键。这不仅表现为对已编入法律和治理之中的伦理所进行的显然可见的争夺（见下文对Floridi区分硬伦理与软伦理的讨论），Facebook、Google及其他行动者大量规避法律和违法的案例便是例证。这也体现在企业对伦理领域本身的圈占：建立AI伦理小组，介入学术界和行动者的争论（例如Facebook的FAIR小组或微软的Aether委员会），并制定判断和治理AI与数据使用的标准本身。

正如我们所论证的，试图界定人工智能伦理包含什么，表明了伦理在争夺AI领导地位的国际竞争中所取得的作用。ISO 31000及其从澳大利亚地方标准发展为其模型被复制和援引的全球标准的历程，正是一个很好的例子：它不仅展示了影响风险和伦理相关关键概念的界定所涉及的地缘政治利害，也说明标准这样的基础设施可以交替作为国家之外的治理术来运作，或推进特定民族国家的利益。

风险基础设施与AI伦理的地缘政治

ISO 31000的案例融入了澳大利亚一种特定的国家想象与地缘政治抱负。澳大利亚风险管理标准AS/NZS 4360:2004所提出的风险概念被ISO 31000采纳后，便成为随后规范人工智能使用的文件文本解释风险的基础，尤其体现在风险管理标准ISO/IEC 23894和AI可信赖性标准之中。^[24]这种联系成为澳大利亚国家宣称自身参与塑造人工智能全球监管框架的某种依据，也提供了一个典范模型，说明如何再次复制并确立这种影响力。澳大利亚标准协会本身是ISO/IEC JTC 1/SC 42工作组的成员，此外还建立了本地对口委员会IT-043，复制国际委员会的工作，并将制定完成的标准引入澳大利亚。本地委员会的构成与澳大利亚标准协会的声明，使我们有理由认为，在塑造AI伦理方面建立鲜明形象，是澳大利亚参与国际委员会的主要目标之一。对口委员会主席Aurelie Jacquet深入参与AI伦理与可信赖性的工作，便体现了这些抱负。正如她所解释的，伦理以及伦理规范和规制的建立，是她所在小组任务的核心部分。^[25]

标准新近向伦理领域延伸，或许值得欢迎，因为这承认了众多批判学者和行动者为凸显AI中的伦理考量所作的工作。然而，ISO 31000及其周围的抱负揭示，技术区中国家之外的治理术空间如何会被地方性的国家愿望所浸染。通过这些纠缠，风险被操作化，并被纳入关于AI及其新兴伦理框架的新的地缘政治构型与经济目标之中。在一些最新文件中，澳大利亚标准制定组织——澳大利亚标准协会

(Standards Australia, SA)——为该国在AI创新领域争取领导地位勾画了一条十分具体的路径。在2020年的报告《人工智能标准路线图：让澳大利亚的声音被听见》中，SA主张，澳大利亚可以采取另一条路径，通过引领标准的制定与实施，尤其是风险与伦理相关标准，确保自身在新兴经济中的领先地位。该文件以国际市场的语言讨论这一可能性，指出澳大利亚不具备输出AI技术的条件，而技术出口是在全球AI竞赛中获得影响力与声望的更显而易见的路径。思考本国AI产业这一被认为存在的不足时，SA把标准和指导原则的输出视为替代性的经济与政治战略。它尤其将ISO 31000视为未来可以复制的案例：一个本土标准被输出，并成为具有影响力的全球标准。援引ISO 31000模型，意味着在制定AI战略和标准时，风险不仅作为一个概念被操作化，也通过利用风险管理框架获利的现有模型被操作化。风险管理标准被视为应当复制的模型——既是一个具有可出口性的案例，更是一个重新思考民族国家在标准制定的国际政治中所扮演角色和功能的案例。

最后这一点是SA审议的重要部分。输出标准的想法，通过把澳大利亚构想为新标准开发试验场的可能性来表述：

国际标准继续为以一致的方式设计、开发、部署和评估AI提供最佳渠道。不过，鉴于学术界、咨询业和一些企业正在积极提出、开发和试验AI系统的风险管理方法与审计方法，我们有机会将其中一些经验编纂起来，形成能够证明澳大利亚专业知识、经验和可行解决方案的文件。这随后可能成为一项国际标准的基础。此事已有先例：澳大利亚利益相关者曾在AS/NZS 4360（风险管理）的制定中发挥重要作用，该标准后来经过完善，被采纳为国际标准（ISO 31000:2009，风险管理——原则与指南）。实现这一目标的最佳方式，是在澳大利亚标准协会内部设立一个汇集不同专长的专门中心。它可以提供类似NIST路线图所提及的试验场，让可能构成标准内容基础的具体提案，与产业及其他利益相关者共同接受检验。^[26]

SA设想把澳大利亚定位为伦理和风险管理等领域全球AI标准开发的试验场或实验中心，由此提出了一种本身就在多个层面操作化风险的模型。将国家视为试验场的想法本身，充满实验与风险之间充满争议的关系。Melinda Cooper^[27]在关于临床试验和实验劳动的研究中指出，后福特主义资本主义拥抱了一种新的风险政治经济学，尤其在信息技术与生物医学领域，它把风险重新界定为价值来源。这种对待风险的新方式，将实验、创新与预期由劳动者和实验对象承担风险所释放的剩余联系起来。在澳大利亚这样一个监管相对严格的国家，就卫生与医疗领域而言，Cooper的结论似乎相关性较弱。然而，在数据提取等新兴经济领域，该国保护水平相对较低而数字技术采用率较高，这种组合使其成为理想的实验场所。我们很容易看出，将澳大利亚建成标准开发试验场的抱负，以及SA于2004年引入的风险偏好概念，都与晚期资本主义中这种新的风险与实验经济相吻合。这种经济具有鲜明的地缘政治轮廓，使资本得以利用监管环境、数据积累和公开宣称的“风险偏好”之间某些特定的理想交汇点。

不过，SA的提议也通过将风险重新概念化为治理技术，引入了对国家如何与其主体发生关系的新理解。具体而言，这种重新界定通过“创新试验场”的概念实现；在信息技术产业中，这一观念早已通过州政府支持的各类创业平台、创新中心及其他实验区得到广泛接受。把国家视为实验空间，延续了风险与控制相互纠缠的历史遗产；这种纠缠正处于澳大利亚作为定居殖民国家和流放殖民地的建国基础之中^[28]，也呼应了2000年代和2010年代在边境管制与强制拘留方面开展的其他“激进”政策实验。这些实验被灾难性地输出到欧洲及其他地区，而人的流动在这些地区日益受到监控和军事化管制。事实上，风险管理的标准化产生了一种可推广的政治经济计算模型，它既能跨越不同机构类型（国家、企业、超国家或其他“国家之外”的机构），也能跨越不同计算对象（人类主体以及与其相关的电子数据）。在地缘政治层面，此处重要的是中等强国的作用：它们试图通过履行制定技术标准等

特定关键功能，将自己嵌入AI超级大国的竞争动态之中；在这些高风险运作中，此类功能同时还能被描绘为善意且政治中立的。

在这一背景下，伦理作为一种复杂的三元市场装置的一部分，取得了特定作用；这一装置将伦理与风险和价值联系起来。在政策文件和标准制定工作中，监管AI并通过可信赖性和伦理等概念施加一定程度伦理监督的必要性，与在全球AI经济空间争取领导地位的抱负紧密相连。2019年，美国商务部国家标准与技术研究院宣称：

“美国在AI领域的全球领导地位，取决于联邦政府在AI标准制定中发挥积极且目标明确的作用。”^[29]澳大利亚标准协会2020年的报告呼应了标准建构与AI领导地位之间的纠缠，其标题《人工智能标准路线图：让澳大利亚的声音被听见》强化了标准能够帮助确立本国在人工智能领域领导地位的观念。欧盟《数字战略》表达了类似抱负：拟议的AI法律框架被视为一种让“欧洲能够在全球发挥领导作用”的手段。^[30]中国的《新一代人工智能伦理规范》（国家新一代人工智能治理专业委员会，2021）虽然没有直接使用同样的表述，但在西方分析中，也大多被解读为中国政府在机器学习和人工智能领域谋求领导地位的表现。^[31]

这些文件表达了国家和区域在全球版图中引领创新的抱负；这一版图沿着机器学习与自动化仍在形成的轮廓展开。同时，这些框架通过定义、标准化与操作化，也构造出被暗示为普遍适用的AI伦理参数，以界定并对冲沿这些创新前沿凝结的风险。这两种目的与其说相互矛盾，不如说组织起了伦理与风险之间的一种特定关系。而且，它们有助于解释，为何澳大利亚这样的中等强国偶尔会在标准制定安排中如此突出：它们相对不受瞩目，充当试验场或实验室，成功经验可以通过同中国、美国和欧盟等具有重大技术利益的国家和地区进行谈判而扩大规模。支持此类实验的国家，可以从直接的“突破性”技术（例如澳大利亚的WiFi），以及在技术创新周期中暂时摆脱边缘地位中获益。

然而，伦理AI以及总体上AI的标准制定安排，多半仍在研发中心附近展开。数据经济中的领先地位，与对于何谓伦理AI的具体参数的建构之间的联系，大体上并不抑制企业通过挖掘数据和训练机器学习模型谋利。相反，它们明确了一组坐标，尤其要求企业行动者在其中行进。民族国家也并非仅仅给标准盖章：中美政府对阿里巴巴、腾讯、Facebook、Google、Apple和Microsoft等科技公司进行监督和质询的情形尽管差异很大，却共同表明，当监管行动者与企业行动者在我们所称的“风险—伦理—价值”三元方程的计算上出现分歧时，这些坐标能够被增设、重新定位或重新强调。

Roxana Radu^[32]指出：“拥有科技产业巨头的国家已经取得领先，抱负是在未来十年中主导全球AI发展。”通过监管谋求AI市场领导地位的抱负充满悖论与冲突，揭示了国家政府与跨国科技巨头之间不稳固的相互依赖。那些怀抱全球抱负的地方倡议，尤其在开发可信赖、公平、负责任或合乎伦理的AI领域，其目标并不总是清楚的：究竟是市场监管、政府监督，还是通过技术官僚手段实现地缘经济支配？国

内的动员也不能被忽略，包括对技术越界的批评；在美国，政治左翼和右翼都发出了这种批评，即使这种动员的结果似乎可能限制民族主义抱负。监管呼声的另一个悖论，是最可能受其影响的企业谨慎地表示支持。例如，Facebook首席执行官Mark Zuckerberg赞成政府加强对其公司运营的监督；他无疑知道，市场既有巨头更容易承担监管合规的成本，而且可以通过向标准机构投入资源及政治游说来控制这些成本。这种支持背后的迂回逻辑，颠倒了Birnhack和Elkin-Koren^[33]所描述的早期互联网公私关系逻辑：当时国家动员在其辖区内注册的私营企业，监管和治理刚刚兴起的网络空间。Facebook呼吁监管，是在不合规风险（如罚款）相对容易承担、且一旦适当立法出台便可由预期提高的公众信任加以抵消的情境中，要求一种互惠形式的保护。正如学者所论述的，大型科技企业伦理指导原则与规范的支持表明，伦理已经被工具化，用来规避——或同样可能用来引导——政府的监督与监管。^[34]

欧盟《通用数据保护条例》（GDPR）已经证明，以数据与数字技术为重点的立法工具和标准化尝试，能够产生跨境影响。^[35] GDPR之所以能够影响世界各地的公司，是因为欧盟立法的特定适用范围——它适用于欧盟公民的数据——以及在平台和网络服务的覆盖与使用都遍及全球的情况下，为每位用户实施不同规则的困难，较小组织尤其如此。GDPR模型表明，针对数字技术与创新监管的地方倡议，能够塑造数据产业的全球地缘政治和地缘经济格局，将国家或超国家行动者确立为领导者；这不仅具有象征意义，也对数字经济产生实际影响。Metzinger^[36]作为AI伦理指导原则制定的“内部人”指出，美国、中国这样的单一民族国家立场矛盾：它们一方面监管企业，另一方面又与具有强烈国家关联的企业站在一起提供支持；这实际上意味着，超国家或联邦式团体有责任制定这样的指导原则，让世界其他地区遵循。

与此同时，GDPR及其对数字创新的影响，也使当下环境中“何谓地缘政治”这一概念复杂化。欧盟立法在重构政治经济权力方面产生了一些相互矛盾的后果。一方面，人们普遍认为，它确立了欧盟及所谓“欧洲价值观”对数据驱动创新未来发展的影响。^[37] 另一方面，作为上文所说监管悖论效应的具体案例，GDPR因增加隐私合规监测负担，将一些小型参与者挤出了竞争，从而意外地巩固了数字空间中的企业权力与垄断。^[38] Johnson、Shiver和Goldberg指出，GDPR出台并获采纳后不久，便导致Google和Facebook等科技巨头的网络合作伙伴数量下降，而这些大公司的市场份额却上升。^[39] 这个由GDPR引发、或许并非有意为之的市场集中案例，凸显了参与数据监管制定、并受到国家、区域和跨国倡议影响的众多行动者之间关系的复杂性。

在中国的案例中，地缘政治行动者的角色也存在类似的模糊性。中国对于人工智能以及数据使用和机器学习伦理的监管倡议，常被通过一个整齐划一的一党制国家的观念来解读。然而，行动者的构成更加多样。中国的AI发展战略，涉及中央和地方

政府以及阿里巴巴、百度、华为等选定的“国家龙头”企业之间的协调。^[40]值得注意的是，中国政府于2021年末发布的《新一代人工智能伦理规范》^[41]，将市场竞争规则纳入合乎伦理的AI生产与供给之中。这些规则可以被作不同解释，但无论如何都表明，对伦理的理解建立在混合的政治与经济议程之上。尽管这一举措及中国政府其他行动的影响，已导致阿里巴巴等公司的市值萎缩^[42]，但它们的似乎既是使企业利益与国家利益保持一致——例如中国半导体供应的“自给自足”，而阿里巴巴近期在这一领域取得了令人意外的进展^[43]——也是增强竞争或追求“共同富裕”。^[44]

各式各样的监管案例，试图确立同样多样的伦理数据治理观念。Luciano Floridi^[45]确实提出，在分析数据和AI伦理时，需要区分两个伦理概念：硬伦理（或规范伦理）和软伦理（或合规后伦理）。软伦理在现行立法的参数，以及遵循法律与道德行动规范的可行性范围内运作。它包含计算与妥协，并公开地受到既有政治经济现实的驱动和约束。Floridi将软伦理与治理体系的渐进发展联系起来，并将欧盟置于那些可以在不损害人权的情况下应用软伦理的政治实体的前列；尽管如此，他对作为道德哲学的伦理，以及立法和技术官僚式合规、监管规范语境下的伦理所作的区分，揭示了我们所据以展开论述的数据伦理与AI伦理的一个重要方面。数字技术与AI领域的伦理，正越来越多地被产业监管和自我监管倡议重新塑造和界定。这一发展指向伦理概念得以建构的充满争议的政治领域；它也塑造了西方道德哲学中那些常被视为普遍适用、且未受质疑的规范性概念。实际上，Floridi对GDPR的高度评价，可以被视为推进了一种貌似普遍主义的道德观，却忽视了原住民数据主权等根植地方经验和政治基础的数据伦理原则的兴起。道德哲学的欧洲中心主义已经受到多种立场的批评；Rosi Braidotti、Nikita Dhawan和Homi Bhabha等作者^[46]，从女性主义、后人类主义和后殖民研究的视角，质疑其普遍适用性的假设。悖论在于，产业主导的数据伦理概念的出现，再次提醒我们：确立一个伦理实践领域及界定该领域的范畴，本身就是政治工作。在ISO 31000的案例中，我们看到伦理概念是在与政治经济治理另外两个关键概念——风险和价值——的关系中被构造出来的。

结论

风险与伦理在标准制定和立法文件的社会技术基础设施中被操作化，表明国家、公司和超国家组织正在行进于、也在建构一种新的地缘政治框架；这一框架规定伦理、风险，以及通过控制和管理装置对二者进行中介意味着什么。我们认为，价值与另外这些术语之间存在某种悖论关系。一方面，价值动摇了任何以标准为支撑的地缘政治确定性，产生风险运作的新向量，并对普遍伦理原则的可能性提出质疑。另一方面，价值的不同规定——对企业而言是经济价值，对国家而言是地缘政治价值，而至少在专注于AI益处的文献中，对那些其数据可能在医疗、法律或消费机构手中被加工为更准确预测的人而言，则是认识论价值——也为承担风险、制定伦理原则准备了基础。这三个概念及其不断变动的构型，一同帮助组织数据交换和算法生产的市场。

在AI监管语境中，这种相互关系通过两种不同的风险概念化方式运作。第一，在风险管理、先发控制和画像分析的算法与工具中，AI被表述为消除风险的技术。^[47]第二，通过“风险偏好”概念，AI进入了有关个人、社群和民族国家可容忍风险程度的模糊关系中；这些风险被辩护为追求机器学习与AI领域领导地位和创新的一部分。由数据支持的决策和系统所生成的伤害风险，作为这一新经济中的生产性行动因素，推动风险本身的资本化；这一举措虽在实践上有所不同，却与金融领域对风险的操作化一致。^[48]这种操作化似乎是组织处理大数据时的财务和社会行为框架的一部分；在此意义上，每一次数据泄露或数据伤害事件，都有助于为这种操作化创造新的市场。应对偶然性的策略，越来越多地以“风险偏好”声明为模型^[49]；在数据经济中，这些声明优先考虑共享与互操作性，以释放数据集的价值潜力，同时也计入风险管理所固有的成本。

大数据价值的复杂构成，重申了Simmel所描述的新经济社会关系中信任与风险之间的依赖关系。伦理在此并非一种置于这些关系之外、并凌驾于其上的抽象道德范畴或准则，而是提供社交性与信任的框架，使大数据价值得以在其中生产和流通。澳大利亚标准协会及其通过标准化谋求全球AI领导地位的抱负表明，风险的社会化正在成为数据经济的核心部分，不仅发生在生产与使用领域，也发生在监管领域。这造成风险与伦理之间的一种悖论关系：AI和数据伦理规制的发展，提供了控制数据所生风险的框架；但与此同时，伦理本身又成为通过实验中心和实验室实践使新风险形式社会化的载体，而这些实践能够扩大至整个民族国家的层级。这一论断并未减损社会压力和声誉利害在推动大数据利用采纳伦理实践方面的重要性^[50]，反而凸显了其重要性；它强调，对于我们如今共同生产的海量社会数据集这类高度社会化的资源，信任与伦理的管理，已经成为从中提取经济价值不可分割的一部分。

资助披露

本文工作属于“自动化的地缘政治”项目（The Geopolitics of Automation, ID: GA64648）的一部分，由澳大利亚研究理事会“发现”资助计划提供资助。

参考文献

- Akhigbe, A., Martin, A.D., and Whyte, A.M. ‘Dodd-Frank and Risk in the Financial Services Industry’, *Review of Quantitative Finance and Accounting*, 47(2) (2016): 395–415. 中文题名：《〈多德—弗兰克法案〉与金融服务业的风险》。
- Amoore, L. ‘Data Derivatives: On the Emergence of a Security Risk Calculus for our Times’, *Theory, Culture & Society*, 28(6), (2011): 24–43. 中文题名：《数据衍生物：论我们时代安全风险计算的兴起》。
- . ‘Security and the Incalculable’, *Security Dialogue*, 45(5), (2014): 423–39. 中文题名：《安全与不可计算之物》。
- . ‘Doubt and the Algorithm: On the Partial Accounts of Machine Learning’, *Theory, Culture & Society*, 36(6), (2019): 147–69. 中文题名：《怀疑与算法：论机器学习的片面说明》。
- . *Cloud Ethics*. Durham: Duke University Press, 2020. 中文题名：《云伦理》。
- Aradau, C. and Blanke, T. ‘The (Big) Data-security assemblage: Knowledge and critique’, *Big Data & Society*, issue 2(2) (2015): 1–12. 中文题名：《（大）数据—安全组装体：知识与批判》。
- . ‘Politics of Prediction: Security and the Time/Space of Governmentality in the Age of Big Data’, *European Journal of Social Theory*, 20(3), (2017): 373–91. 中文题名：《预测的政治：大数据时代的安全与治理术的时间/空间》。
- Barry, A. ‘Technological Zones’, *European Journal of Social Theory*, 9(2), (2006): 239–53. 中文题名：《技术区》。
- Beck, U. *Risk Society: Towards a New Modernity*, Ritter, M. 译, Newbury Park, CA: SAGE Publications, 1992. 中文题名：《风险社会：迈向一种新的现代性》。
- Bhabha, H.K. ‘Culture’ s In-between’, in Hall, S. & du Gay, P. (eds) *Questions of Cultural Identity*, [CA: SAGE Publications], 1996, pp. 53–60. 中文题名：《文化的中间地带》；《文化认同问题》。
- Birnback, M. and Elkin-Koren, N. ‘The Invisible Handshake: The Reemergence of the State in the Digital Environment’, *SSRN Electronic Journal* 8(6), (2003): 1–57. 中文题名：《看不见的握手：国家在数字环境中的重新出现》。
- Brouwer, M.T. ‘Weber, Schumpeter and Knight on Entrepreneurship and Economic Development’, *Journal of Evolutionary Economics*, 12(1), (2002): 83–105. 中文题名：《韦伯、熊彼特与奈特论企业家精神和经济发展》。
- Braidotti, R. *Transpositions: On Nomadic Ethics*. Cambridge: Polity, 2006. 中文题名：《转位：论游牧伦理》。

- Clarke, R. 'Big Data, Big Risks', *Information Systems Journal*, 26(1), (2016): 77–90. 中文题名: 《大数据, 大风险》。
- Cooper, M. 'Experimental Labour—Offshoring Clinical Trials to China', *East Asian Science, Technology and Society: An International Journal*, 2(1), (2008): 73–92. 中文题名: 《实验劳动: 将临床试验离岸外包至中国》。
- Cooper, M.E. *Life as Surplus: Biotechnology and Capitalism in the Neoliberal Era*. Seattle: University of Washington Press, 2011. 中文题名: 《作为剩余的生命: 新自由主义时代的生物技术与资本主义》。
- Daly, A. 'Neo-liberal Business-as-Usual or Post-Surveillance Capitalism with European Characteristics? The EU's General Data Protection Regulation in a Multi-Polar Internet', in Hoyng, R. & Pak Lei Chong, G. (eds), *Critiquing Communication Innovation New Media in a Multipolar World. US-China Relations in the Age of Globalization*, Michigan State University Press, 2022, pp. 29–54. 中文题名: 《新自由主义的一切照旧, 还是具有欧洲特色的后监控资本主义? 多极互联网中的欧盟〈通用数据保护条例〉》; 《批判传播创新: 多极世界的新媒体; 全球化时代的美中关系》。
- Deloitte. 'Risk Appetite Frameworks: How to Spot the Genuine Article', 可见: <https://www2.deloitte.com/content/dam/Deloitte/au/Documents/risk/deloitte-au-risk-appetite-frameworks-financial-services-0614.pdf>, 2014. 中文题名: 《风险偏好框架: 如何辨认真货》。
- Dhawan, N. 'Can Non - Europeans Philosophize? Transnational Literacy and Planetary Ethics in a Global Age', *Hypatia*, 32(3), (2017): 488–505. 中文题名: 《非欧洲人能进行哲学思考吗? 全球时代的跨国素养与行星伦理》。
- Easterling, K. *Extrastatecraft: The Power of Infrastructure Space*. London: Verso Books, 2014. 中文题名: 《国家之外的治理术: 基础设施空间的力量》。
- The National New Generation Artificial Intelligence Governance Specialist Committee. 'Ethical Norms for New Generation Artificial Intelligence', 2021, 英文译本见: https://cset.georgetown.edu/wp-content/uploads/t0400_AI_ethical_norms_EN.pdf. 中文题名: 《新一代人工智能伦理规范》。
- Eubanks, V. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*, New York: St. Martin's Press, 2018. 中文题名: 《自动化不平等: 高科技工具如何为穷人画像、监控和惩罚穷人》。
- European Commission. 'Regulatory Framework Proposal on Artificial Intelligence', 2022, 可见: <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai#:~:text=The%20proposed%20AI%20regulation%20ensures,address%20to%20avoid%20undesirable%20outcomes>. 中文题名: 《人工智能监管框架提案》。
- Galloway, A. *Protocol: How Control Exists after Decentralization*, Cambridge, Massachusetts: MIT Press, 2004. 中文题名: 《协议: 去中心化之后控制如何存在》。
- Goddard, M. 'The EU General Data Protection Regulation (GDPR): European Regulation that has a Global Impact', *International Journal of Market Research* 59 (6, 2017): 703–5. 中文题名: 《欧盟〈通用数据保护条例〉(GDPR): 具有全球影响的欧洲规制》。

- Gstrein, O.J. and Zwitter, A.J. ‘Extraterritorial Application of the GDPR: Promoting European Values or Power?’ , *Internet Policy Review* [在线] 10(3) (2021), 可见: <https://policyreview.info/articles/analysis/extraterritorial-application-gdpr-promoting-european-values-or-power>. 中文题名: 《GDPR的域外适用: 促进欧洲价值观, 还是欧洲权力?》。
- Günther, W.A., Mehrizi, M.H.R., Huysman, M. and Feldberg, F. ‘Debating Big Data: A Literature Review on Realizing Value from Big Data’ , *The Journal of Strategic Information Systems*, 26(3), (2017): 191–209. 中文题名: 《争论大数据: 关于实现大数据价值的文献综述》。
- Floridi, L. ‘Soft Ethics and the Governance of the Digital’ , *Philosophy & Technology*, 31 (1), (2018): 1–8. 中文题名: 《软伦理与数字治理》。
- Foucault, M. *Security, Territory, Population: Lectures at the Collège De France 1977–1978*, G. Burchell 译。New York: Picador, 2007. 中文题名: 《安全、领土、人口: 法兰西学院 1977–1978年讲座》。
- Hoelzl, I. and Marie, R. *Softimage: Towards a New Theory of the Digital Image*, Bristol, UK: Intellect Books, 2015. 中文题名: 《软图像: 迈向数字图像的新理论》。
- Hristova, T. *Data Infrastructures and Digital Labour: The Case of Teleradiology*, 博士学位论文, Western Sydney University, Sydney, 2020. 中文题名: 《数据基础设施与数字劳动: 远程放射学案例》。
- Hristova, T., Neilson, B. and Rossiter, N. ‘Digital Infrastructure, Liminality, and World-Making Via Asia| On the Block Train: Rethinking Block Technologies on the YuXinOu Express’ , *International Journal of Communication* 15 (2021): 2613–2630. 中文题名: 《数字基础设施、阈限性与经由亚洲的世界建构 | 集装箱班列: 重新思考渝新欧快线上的区块链技术》。
- International Research Center for AI Ethics and Governance. *The Ethical Norms for the New Generation Artificial Intelligence, China*, 2021, 可见: <https://ai-ethics-and-governance.institute/2021/09/27/the-ethical-norms-for-the-new-generation-artificial-intelligence-china/>. 中文题名: 《中国〈新一代人工智能伦理规范〉》。
- ISO/IEC. 2022. International standard ISO/IEC DIS 23894. Information technology—Artificial intelligence—Risk management. International Organization for Standardization/International Electrotechnical Commission, (草案), 可见: <https://www.iso.org/standard/77304.html>. 中文题名: 《国际标准ISO/IEC DIS 23894: 信息技术——人工智能——风险管理》。
- ISO/IEC 27001. Information security management. International Organization for Standardization/International Electrotechnical Commission, 2013, <https://www.iso.org/standard/54534.html>. 中文题名: 《信息安全管理》。
- ISO Guide 73:2009. Risk management—Vocabulary, International Organization for Standardization, 2009, <https://www.iso.org/obp/ui/#iso:std:iso:guide:73:ed-1:v1:en>. 中文题名: 《风险管理——词汇》。

- ISO 31000: 2018 Risk management—Guidelines, International Organization for Standardization, 2018, <https://www.iso.org/obp/ui/#iso:std:iso:31000:ed-2:v1:en>. 中文题名：《风险管理——指南》。
- Johnson, G., Shriver, S. and Goldberg, S. ‘Privacy & Market Concentration: Intended & Unintended Consequences of the GDPR’ , 2021, <https://dx.doi.org/10.2139/ssrn.3477686>. 中文题名：《隐私与市场集中：GDPR的预期及非预期后果》。
- Kitchin, R. ‘Thinking Critically About and Researching Algorithms’ , *Information, Communication & Society*, 20(1), (2017): 14–29. 中文题名：《批判性地思考与研究算法》。
- LeCun, Y. ‘Deep Learning Hardware: Past, Present, and Future’ , *IEEE International Solid-State Circuits Conference-(ISSCC)* 2019: 12–19. 中文题名：《深度学习硬件：过去、现在与未来》。
- Metzinger, T. ‘Ethics Washing Made in Europe’ , *Der Tagesspiegel*, 8 April 2019, <https://www.tagesspiegel.de/politik/eu-guidelines-ethics-washing-made-in-europe/24195496.html>. 中文题名：《欧洲制造的伦理洗白》。
- Munn, L., Hristova, T., & Magee, L. ‘Clouded Data: Privacy and the Promise of Encryption’ , *Big Data & Society*, 6(1), (2019): 1–16. 中文题名：《云雾中的数据：隐私与加密的承诺》。
- Murphy, C.N. and Yates, J. *The International Organization for Standardization (ISO): Global Governance through Voluntary Consensus*. Milton Park: Routledge, 2009. 中文题名：《国际标准化组织（ISO）：通过自愿共识进行全球治理》。
- Naden, C. ‘It’ s All About Trust’ , 2019, <https://www.iso.org/news/ref2452.html>. 中文题名：《一切关乎信任》。
- National Institute for Standards and Technology. ‘U.S. Leadership in AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools Prepared in Response to Executive Order 13859’ , 9 August 2019, https://www.nist.gov/system/files/documents/2019/08/10/ai_standards_fedengagement_plan_9aug2019.pdf. 中文题名：《美国在AI领域的领导地位：响应第13859号行政令制定的联邦参与技术标准与相关工具开发计划》。
- Niebel, C. ‘The Impact of the General Data Protection Regulation on Innovation and the Global Political Economy’ , *Computer Law & Security Review*, 40, (2021) 105523: p. 1–15. 中文题名：《〈通用数据保护条例〉对创新和全球政治经济的影响》。
- Power, M. ‘The Risk Management of Everything’ , *The Journal of Risk Finance*, 5(3), (2004): 58–65. 中文题名：《一切事物的风险管理》。
- Productivity Commission. ‘Data Availability and Use: Productivity Commission Inquiry Report’ , 2017. 可见：<https://www.pc.gov.au/inquiries/completed/data-access/report/data-access.pdf>. 中文题名：《数据可得性与使用：生产力委员会调查报告》。
- Radu, R. ‘Steering the Governance of Artificial Intelligence: National Strategies in Perspective’ , *Policy and Society* 40(2), (2021): 178–93. 中文题名：《引导人工智能治理：国家战略透视》。

- Roberts, H., Cows, J., Morley, J. et al. 'The Chinese Approach to Artificial Intelligence: An Analysis of Policy, Ethics, and Regulation', *AI & Soc* 36, (2021): 59–77. 中文题名：《中国的人工智能路径：政策、伦理与监管分析》。
- Silverpond. 'The Role of Ethics in AI Development, Implementation and Governance', 2021, <https://silverpond.com.au/ai-community/australian-ai-ecosystem-survey/aurelie-jacquet-2020-2021-australian-ai-ecosystem-survey/>. 中文题名：《伦理在AI开发、实施与治理中的作用》。
- Simmel, G. *The Philosophy of Money*, Milton Park: Routledge, 2004. 中文题名：《货币哲学》。
- Standards Australia. *An Artificial Intelligence Standards Roadmap: Making Australia's Voice Heard. Final Report*, 2020, <https://www.standards.org.au/getmedia/ede81912-55a2-4d8e-849f-9844993c3b9d/1515-An-Artificial-Intelligence-Standards-Roadmap12-02-2020.pdf.aspx>. 中文题名：《人工智能标准路线图：让澳大利亚的声音被听见。最终报告》。
- Veracini, L. 'Understanding Colonialism and Settler Colonialism as Distinct Formations', *Interventions*, 16(5), (2014): 615–33. 中文题名：《将殖民主义与定居殖民主义理解为不同的构成形态》。
- Wagner, B. 'Ethics as an Escape from Regulation. From "Ethics-washing" to Ethics-shopping?' in Bayamlioglu, E., Baraliuc, I., Janssens, L.A.W., Hildebrandt, M. (eds) *Being Profiled*, Amsterdam: Amsterdam University Press, 2018, 84–9. 中文题名：《作为逃避监管手段的伦理：从“伦理洗白”到挑选伦理？》；《被画像》。
- Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. 'The FAIR Guiding Principles for Scientific Data Management and Stewardship', *Sci Data* 3, 160018, (2016): 1–9. 中文题名：《科学数据管理与托管的FAIR指导原则》。
- Yahoo Finance. Alibaba Group Holding Limited (BABA), 2022, <https://finance.yahoo.com/quote/BABA/>. 中文题名：《阿里巴巴集团控股有限公司（BABA）》。

原注

- [1] M. Wilkinson, M. Dumontier, I. Aalbersberg, et al., 'The FAIR Guiding Principles for Scientific Data Management and Stewardship', *Sci Data* 3, 160018, (2016). 中文题名：《科学数据管理与托管的FAIR指导原则》。返回
- [2] C.N. Murphy, and J. Yates, *The International Organization for Standardization (ISO): Global Governance through Voluntary Consensus*. Milton Park: Routledge, 2009. 中文题名：《国际标准化组织（ISO）：通过自愿共识进行全球治理》。返回
- [3] ISO/IEC, 'Draft international standard ISO/IEC DIS 23894', Information technology—Artificial intelligence—Risk management, 2022. 中文题名：《国际标准草案ISO/IEC DIS 23894》；《信息技术——人工智能——风险管理》。返回
- [4] Keller Easterling, K. *Extrastatecraft: The Power of Infrastructure Space*. London: Verso Books, 2014. 中文题名：《国家之外的治理术：基础设施空间的力量》。返回

- [5] Andrew Barry, 'Technological Zones', *European Journal of Social Theory*, 9(2), (2006): 239–53. 中文题名：《技术区》。返回
- [6] C. Aradau and T. Blanke, 'The (Big) Data-security assemblage: Knowledge and critique', *Big Data & Society*, 2(2) (2015): 1–12; R. Kitchin, 'Thinking Critically About and Researching Algorithms', *Information, Communication & Society*, 20(1), (2017): 14–29. 中文题名：《（大）数据—安全组装体：知识与批判》；《批判性地思考与研究算法》。返回
- [7] Alexander Galloway. *Protocol: How control exists after decentralization*. MIT Press. 2004. 中文题名：《协议：去中心化之后控制如何存在》。返回
- [8] I. Hoelzl and R. Marie, *Softimage: Towards a New Theory of the Digital Image*, Bristol, UK: Intellect Books, 2015. 中文题名：《软图像：迈向数字图像的新理论》。返回
- [9] T. Hristova, *Data Infrastructures and Digital Labour: The Case of Teleradiology*, 博士学位论文, Western Sydney University, Sydney, 2020. 中文题名：《数据基础设施与数字劳动：远程放射学案例》。返回
- [10] Easterling, *Extrastatescraft*. 中文题名：《国家之外的治理术》。返回
- [11] L. Amore, 'Data Derivatives: On the Emergence of a Security Risk Calculus for our Times', *Theory, Culture & Society*, 28(6), (2011):24–43; Amore, 'Security and the Incalculable', *Security Dialogue*, 45(5), (2014): 423–39; Amore, *Cloud Ethics*. Durham: Duke University Press, 2020; C. Aradau and T. Blanke, 'The (Big) Data-security assemblage: Knowledge and critique', *Big Data & Society*, issue 2(2) (2015): 1–12; V. Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*, New York: St. Martin's Press, 2018, 等诸多文献。中文题名：《数据衍生物：论我们时代安全风险计算的兴起》；《安全与不可计算之物》；《云伦理》；《（大）数据—安全组装体：知识与批判》；《自动化不平等：高科技工具如何为穷人画像、监控和惩罚穷人》。返回
- [12] Hristova, *Data Infrastructures and Digital Labour*; T. Hristova, B. Neilson, and N. Rossiter, 'Digital Infrastructure, Liminality, and World-Making Via Asia On the Block Train: Rethinking Block Technologies on the YuXinOu Express', *International Journal of Communication* 15 (2021). 中文题名：《数据基础设施与数字劳动》；《数字基础设施、阈限性与经由亚洲的世界建构：集装箱班列——重新思考渝新欧快线上的区块技术》。返回
- [13] IPCC, '2—Determinants of Risk: Exposure and Vulnerability—IPCC', https://www.ipcc.ch/pdf/special-reports/srex/SREX-Chap2_FINAL.pdf. 中文题名：《2——风险的决定因素：暴露度与脆弱性——IPCC》。返回
- [14] G. Simmel, *The Philosophy of Money*, Milton Park: Routledge, 2004, 177–8. 中文题名：《货币哲学》。返回
- [15] U. Beck, *Risk Society: Towards a New Modernity*, Ritter, M. 译, Newbury Park, CA: SAGE Publications, 1992. 中文题名：《风险社会：迈向一种新的现代性》。返回
- [16] (ISO 31000: 2018), 第3.1条。返回
- [17] (ISO Guide 73:2009), 第3.7.1.2条。返回

- [18] Deloitte, 'Risk Appetite Frameworks: How to Spot the Genuine Article', 可见: <https://www2.deloitte.com/content/dam/Deloitte/au/Documents/risk/deloitte-au-risk-appetite-frameworks-financial-services-0614.pdf>, 2014. 中文题名: 《风险偏好框架: 如何辨认真货》。返回
- [19] Y. LeCun, 'Deep Learning Hardware: Past, Present, and Future', *IEEE International Solid-State Circuits Conference-(ISSCC)* 2019: 12–19. 中文题名: 《深度学习硬件: 过去、现在与未来》。返回
- [20] M. Power, 'The Risk Management of Everything', *The Journal of Risk Finance*, 5(3), (2004): 58–65. 中文题名: 《一切事物的风险管理》。返回
- [21] L. Munn, T. Hristova, and L. Magee, 'Clouded Data: Privacy and the Promise of Encryption', *Big Data & Society*, 6(1), (2019): 1–16. 中文题名: 《云雾中的数据: 隐私与加密的承诺》。返回
- [22] Amoores, 'Doubt and the Algorithm: On the Partial Accounts of Machine Learning', *Theory, Culture & Society*, 36(6), (2019): 147–69. 中文题名: 《怀疑与算法: 论机器学习的片面说明》。返回
- [23] M.T. Brouwer, 'Weber, Schumpeter and Knight on Entrepreneurship and Economic Development', *Journal of Evolutionary Economics*, 12(1), (2002): 83–105. 中文题名: 《韦伯、熊彼特与奈特论企业家精神和经济发展》。返回
- [24] C. Naden, 'It's All About Trust', 2019, <https://www.iso.org/news/ref2452.html>. 中文题名: 《一切关乎信任》。返回
- [25] Silverpond, 'The Role of Ethics in AI Development, Implementation and Governance', 2021, <https://silverpond.com.au/ai-community/australian-ai-ecosystem-survey/aurelie-jacquet-2020-2021-australian-ai-ecosystem-survey/>. 中文题名: 《伦理在AI开发、实施与治理中的作用》。返回
- [26] Standards Australia, *An Artificial Intelligence Standards Roadmap: Making Australia's Voice Heard. Final Report*, 2020, 35–36, <https://www.standards.org.au/getmedia/ede81912-55a2-4d8e-849f-9844993c3b9d/1515-An-Artificial-Intelligence-Standards-Roadmap12-02-2020.pdf.aspx><https://www.standards.org.au/getmedia/ede81912-55a2-4d8e-849f-9844993c3b9d/1515-An-Artificial-Intelligence-Standards-Roadmap12-02-2020.pdf.aspx>. 中文题名: 《人工智能标准路线图: 让澳大利亚的声音被听见。最终报告》。[译注: 原书此处网址重复拼接, 照录; 可参照本章参考文献中同一报告的完整网址。] 返回
- [27] M. Cooper, 'Experimental Labour—Offshoring Clinical Trials to China', *East Asian Science, Technology and Society: An International Journal*, 2(1), (2008): 73–92; Cooper, M.E. *Life as Surplus: Biotechnology and Capitalism in the Neoliberal Era*. Seattle: University of Washington Press, 2011. 中文题名: 《实验劳动: 将临床试验离岸外包至中国》; 《作为剩余的生命: 新自由主义时代的生物技术与资本主义》。返回
- [28] L. Veracini, 'Understanding Colonialism and Settler Colonialism as Distinct Formations', *Interventions*, 16(5), (2014): 615–33. 中文题名: 《将殖民主义与定居殖民主义理解为不同的构成形态》。返回

- [29] National Institute for Standards and Technology, ‘U.S. Leadership in AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools Prepared in Response to Executive Order 13859’, 9 August 2019. 中文题名：《美国在AI领域的领导地位：响应第13859号行政令制定的联邦参与技术标准与相关工具开发计划》。返回
- [30] European Commission, ‘Regulatory Framework Proposal on Artificial Intelligence’, 2022, <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai#:~:text=The%20proposed%20AI%20regulation%20ensures,address%20to%20avoid%20undesirable%20outcomes>. 中文题名：《人工智能监管框架提案》。返回
- [31] H. Roberts, J., Cows, J. Morley, J. et al. ‘The Chinese Approach to Artificial Intelligence: An Analysis of Policy, Ethics, and Regulation’, *AI & Soc* 36, (2021): 59–77. 中文题名：《中国的人工智能路径：政策、伦理与监管分析》。返回
- [32] R. Radu, ‘Steering the Governance of Artificial Intelligence: National Strategies in Perspective’, *Policy and Society* 40(2), (2021): 182. 中文题名：《引导人工智能治理：国家战略透视》。返回
- [33] M. Birnhack and N. Elkin-Koren, ‘The Invisible Handshake: The Reemergence of the State in the Digital Environment’, *SSRN Electronic Journal* 8(6), 2003: 1–57 中文题名：《看不见的握手：国家在数字环境中的重新出现》。返回
- [34] T. Metzinger, ‘Ethics Washing Made in Europe’, *Der Tagesspiegel*, 8 April 2019, <https://www.tagesspiegel.de/politik/eu-guidelines-ethics-washing-made-in-europe/24195496.html>; B. Wagner, ‘Ethics as an Escape from Regulation. From “Ethics-washing” to Ethics-shopping?’ in (eds) Bayamlioglu, E., Baraliuc, I., Janssens, L.A. W., Hildebrandt, M., Amsterdam (eds) *Being Profiled*, Amsterdam University Press, 2018, 84–9. 中文题名：《欧洲制造的伦理洗白》；《作为逃避监管手段的伦理：从“伦理洗白”到挑选伦理？》；《被画像》。返回
- [35] M. Goddard, ‘The EU General Data Protection Regulation (GDPR): European Regulation that has a Global Impact’, *International Journal of Market Research*, 59 (6), (2017): 703–5; C. Niebel, ‘The Impact of the General Data Protection Regulation on Innovation and the Global Political Economy’, *Computer Law & Security Review*, 40, (2021): p.105523. 中文题名：《欧盟〈通用数据保护条例〉（GDPR）：具有全球影响的欧洲规制》；《〈通用数据保护条例〉对创新和全球政治经济的影响》。返回
- [36] Metzinger, ‘Ethics Washing Made in Europe’. 中文题名：《欧洲制造的伦理洗白》。返回
- [37] A. Daly, ‘Neo-liberal Business-as-Usual or Post-Surveillance Capitalism with European Characteristics? The EU’s General Data Protection Regulation in a Multi-Polar Internet’, in *Communication Innovation and Infrastructure: A Critique of the New in a Multipolar World*, East Lansing: Michigan State University Press, (即将出版, 2021) pp.66–95; O.J. Gstrein. and A.J. Zwitter, ‘Extraterritorial Application of the GDPR: Promoting European Values or Power?’, *Internet Policy Review* [在线] 10(3) (2021), 可见：<https://policyreview.info/articles/analysis/extraterritorial-application-gdpr-promoting-european-values-or-power>. 中文题名：《新自由主义的一切照旧，还是具有欧洲特色的后监控资本主义？多极互联网中的欧盟〈通用数据保护条例〉》；《传播创

- 新与基础设施：对多极世界中新事物的批判》；《GDPR的域外适用：促进欧洲价值观，还是欧洲权力？》。返回
- [38] G. Johnson, G., S. Shriver, and S. Goldberg, 'Privacy & Market Concentration: Intended & Unintended Consequences of the GDPR', 2021. 3477686. 中文题名：《隐私与市场集中：GDPR的预期及非预期后果》。返回
- [39] Johnson, Shriver, and Goldberg, 'Privacy & Market Concentration: Intended & Unintended Consequences of the GDPR', 15. 中文题名：《隐私与市场集中：GDPR的预期及非预期后果》。返回
- [40] H. Roberts, J., Cows, J. Morley, J. et al., 'The Chinese Approach to Artificial Intelligence: An Analysis of Policy, Ethics, and Regulation'. 中文题名：《中国的人工智能路径：政策、伦理与监管分析》。返回
- [41] International Research Center for AI Ethics and Governance（人工智能伦理与治理国际研究中心），2021。返回
- [42] 截至撰写本文时（6月4日），阿里巴巴在美国上市的控股公司股价已较2020年10月1日304.69美元的高点下跌69%（Yahoo Finance. Alibaba Group Holding Limited (BABA) [阿里巴巴集团控股有限公司]，2022, <https://finance.yahoo.com/quote/BABA/>）。返回
- [43] Dashveenjit Kaur, 'China's most advanced chip may soon come from Alibaba', *Techwire Asia*, 21 October 2021. <https://techwireasia.com/2021/10/the-new-chip-by-alibaba-may-be-one-of-the-most-advanced-in-china/>. 中文题名：《中国最先进的芯片可能很快出自阿里巴巴》。返回
- [44] Brian Liu and Raquel Leslie, 'China's Tech Crackdown: A Year-In-Review', *Lawfare*, 7 January 2022. <https://web.archive.org/web/20220506205610/https://www.lawfareblog.com/chinas-tech-crackdown-year-review>. 中文题名：《中国的科技业整顿：年度回顾》。返回
- [45] L. Floridi, 'Soft Ethics and the Governance of the Digital', *Philosophy & Technology*, 31(1), (2018): 1–8. 中文题名：《软伦理与数字治理》。返回
- [46] R. Braidotti, *Transpositions: On Nomadic Ethics*. Cambridge: Polity, 2006; N. Dhawan, 'Can Non-European Philosophize? Transnational Literacy and Planetary Ethics in a Global Age', *Hypatia*, 32(3), (2017): 488–505; H. Bhabha, 'Culture's In-between', in S. Hall, and P. du Gay (eds) *Questions of Cultural Identity*, London: SAGE, 1996, pp. 53–60. 中文题名：《转位：论游牧伦理》；《非欧洲人能进行哲学思考吗？全球时代的跨国素养与行星伦理》；《文化的中间地带》；《文化认同问题》。返回
- [47] L. Amoore, 'Data Derivatives: On the Emergence of a Security Risk Calculus for our Times'; Amoore, 'Security and the Incalculable'; Amoore, *Cloud Ethics*; 2020, Aradau and Blanke, 'The (Big) Data-security assemblage: Knowledge and critique' Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*, 等诸多文献。中文题名：《数据衍生物：论我们时代安全风险计算的兴起》；《安全与不可计算之物》；《云伦理》；《（大）数据—安全组装体：知识与批判》；《自动化不平等：高科技工具如何为穷人画像、监控和惩罚穷人》。返回

- [48] Cooper, ‘Experimental Labour—Offshoring Clinical Trials to China’ , A. Akhigbe, A.D. Martin, and A.M. Whyte, ‘Dodd–Frank and Risk in the Financial Services Industry’ , *Review of Quantitative Finance and Accounting*, 47(2) (2016): 395–415. 中文题名：《实验劳动：将临床试验离岸外包至中国》；《〈多德—弗兰克法案〉与金融服务业的风险》。[返回](#)
- [49] Productivity Commission（生产力委员会），2017；PricewaterhouseCoopers（普华永道），2012。[返回](#)
- [50] W.A. Günther, M.H.R. Mehrizi, M. Huysman, and F. Feldberg, ‘Debating Big Data: A Literature Review on Realizing Value from Big Data’ , *The Journal of Strategic Information Systems*, 26(3), (2017): 191–209; R. Clarke, ‘Big Data, Big Risks’ , *Information Systems Journal*, 26(1), (2016): 77–90. 中文题名：《争论大数据：关于实现大数据价值的文献综述》；《大数据，大风险》。[返回](#)

军事美德与AI研究中“伦理”的限度

Michael Richardson

战争凭借美德立足：战士的美德、正义事业的美德，以及为维系国家或宗教而在逆境面前做必须做之事的美德。战争的美德至少在西方可以追溯至古代，但其现代表述带有鲜明特征。有德性的战争如今是技术化战争，是凭借后勤、科学与计算的技术革命所赋予的精确性、信息、理性与比例性来实施的战争。在当前转向自主武器系统（AWS）、并更广泛地将人工智能（AI）纳入战争的过程中，美德发挥着说服公众和机构接受日益复杂、日益编码化、日益难以理解的新兴技术之必要性的作用。美德并非只是从自主武器系统和军事AI中流出，而是通过将法律和伦理纳入系统本身而被注入其中。至少，故事是这么说的。正如虚拟性的美德曾帮助遮蔽美国在9/11之后无休止战争中的暴力，拥抱和宣扬自主系统的美德，也有可能遮蔽既有伤害与不公的再生产和加剧，以及新的暴力与压迫形式的形成。

大型科技企业涉足美国军方合同时往往引发争议——想想Google因参与国防部（DoD）的Project Maven计划，把TensorFlow算法用于无人机图像处理而引起的愤怒——但军方与私营及民间机构之间的相互依赖，同样有害地存在于世界各地蓬勃发展的生态之中，其中包括初创企业、研究成果转化中心、国防资助计划、政府倡议、资金拮据的大学和渴求经费的学者。在这些美德经济中，“伦理”不仅通过掩饰军事暴力来促成互利合作，也成为潜在的商品，能够被编码进正在开发的技术本身。^[1]国防研究者和公司不仅能够成为有德性的主体，还能使战争也变得有德性。本章考察澳大利亚新兴军事技术产业及其与学术界的联系，论证军事美德经济的运作方式，与科技产业和大学AI研究之间更广泛联系中的美德经济，既有相似之处，也有差异。

在AI与大数据的开发和专业知识方面，军方远落后于私营部门。五角大楼算法战跨职能小组（AWCFT）的负责人在该小组于2017年成立时承认，这一现实正在加速由产业引领的合作过程，这些过程模仿硅谷敏捷开发模型的某些方面。^[2]鉴于军事技术往往被用于生死攸关的场景，这种“快速行动、打破常规”的方法可能造成严重后果。但军方不仅是大型科技企业的重要客户，也日益成为学术研究重要的资金来源。在军事AI语境中，“伦理”具有一种经济功能，它为产业、学术界与军事机构之间的互动设定框架、提供便利并持续供给动力。军事技术，尤其武器系统，常被置于独特的伦理话语框架之中；这些话语来自武装冲突法、国际人道法、特定且预先界定的交战规则，以及更为模糊的战士精神的混合。包括澳大利亚国防军

（ADF）在内的军方，倾向于将伦理视为工具性的东西，主要涉及士兵个人在战场内外的行为，而不是与正义或社会义务等更大问题相互缠绕。伦理通常被同时表述为应秉持的价值和有待解决的问题。当这一方法遭遇军事AI时，“伦理”作为减少

伤害框架的局限便显现出来：伦理早已从属于军事暴力，因为它始终关心的是如何使这种暴力得以实施。

军事美德

国际关系学者James Der Derian认为，计算在战争中的中心地位在冷战期间开始上升，并在冷战结束后急剧增强，这标志着一种由美国引领的新型武装冲突模式的出现。Der Derian写道：“有德性的战争的核心，是从远距离威胁施加暴力，并在必要时将暴力付诸实现的技术能力和伦理要求——不造成伤亡，或将伤亡降至最低。”^[3]当然，有德性的战争并未消除杀戮，也没有消除对平民的杀戮，美国和以色列使用无人机与自主系统的事实便清楚证明了这一点。有德性的战争也不止于暴力本身：在Der Derian的构想中，战争也是虚拟的，因为它越来越依赖信息与抽象。尽管他在为一种新型战争命名时，试图兼顾“美德”（virtue）与“虚拟”（virtual）两个含义，这种做法有些含混，但他的分析仍指出了两者之间的密切关系：一方面，是依赖模拟、建模、计算、自动化与自主性的新兴战争形式；另一方面，是通过话语重构战争、使其核心地位合法化的做法。正如Der Derian所指出的，有德性的战争由一组边界模糊的机构、行动者与制度生产出来，又反过来维系它们；他将其称为军事—工业—媒体—娱乐网络，简称MIME-NET。^[4]大学及其研究人员就盘旋在这个网络近旁。

无论在美国还是其他地方，军方与学术界之间的紧密联系都远非新事物。冷战期间，国防部（DoD）、中央情报局（CIA）和国防高级研究计划局（DARPA），开始直接资助美国及世界各地众多学科的研究，从核物理、心理学、医学到人类学。DARPA还资助了战略计算计划，该计划在1983年至1993年间向先进计算与人工智能投入了超过10亿美元。不过，军事资助对于计算和控制论本身的出现也至关重要；后者为当代机器学习技术奠定了概念与数学基础。这些技术经常被包装并宣传为“人工智能”^[5]，即使它们在很大程度上依赖人类劳动。^[6]无论应用于军事还是民用环境，当代机器学习技术所依赖的计算能力，常常远远超出大学实验室的承受范围，即使各机构汇集资源也不例外。这促成了Meredith Whittaker所说的大型科技企业对AI研究的“俘获”：研究者依赖于Amazon、Google和Microsoft运营的平台的使用权，因此倾向于开展符合某些AI范式的研究，将既有基础设施不断再现为有益且必要的。^[7]这种依赖关系运作于一系列反复出现的争议的阴影之下，这些争议涉及大型科技企业——尤其Facebook、Google和Amazon——的社会、文化与政治影响。

这些争议反过来加剧了对“美德经济”的需求。在这种经济中，“美德与伦理是主要的生产和流通对象；大型科技企业内部的群体通过设立伦理委员会、工作组等方式生产和传播它们，而大学、研究机构、咨询公司和其他相关产业等外部群体也参与其中。”^[8]正如本书编者最近在一篇关于这一主题的论辩文章中所指出的，这种

经济产生于西方大学公共资助、尤其是研究资助危机日益加深的背景下，而这加强了AI研究对大型科技企业资金等依赖关系。在商业环境中审视这些关系至关重要，但军队、国家安全与情报部门的纠缠，不能被从这个方程中剔除，也不能被当作无关紧要的个案研究。尽管批判性学术研究不断探讨将AI纳入战争和国家安全问题，大学及资助机制在这一过程中的制度角色仍需要得到更多关注。^[9]

Alison Howell令人信服地指出，大学并非最近才被“军事化”，而是始终由尚武政治所生产，并服务于尚武目的机构。^[10]如前所述，更广义的互联网及AI本身，无论直接还是间接，都曾经并仍然在很大程度上依赖军事资助与目标。在当今美国，军事AI开发的重要推动者和国际AI军备竞赛的鼓吹者之一，是Google前首席执行官Eric Schmidt，他如今担任国防技术创新委员会负责人。对Schmidt及其同道而言，军事AI应用从一开始就具有美德，恰恰是因为它们保护国家免受威胁，并强化其在全球舞台上的地位。这一切可以被理解为试图使AI本身安全化：这一领域发生的大量工作，因而进入常规政治之外的领域，并在安全的例外空间中运作。^[11]因此，我们的任务是考察军方如何参与AI研究中的美德经济，因为它们理解伦理并使之商品化的方式，往往不同于民用行动者。

本章通过考察澳大利亚而非美国的情境，处理军方如何参与美德经济的问题。这一考察具有启发性，因为澳大利亚正处于有意识地加速本国军事产业发展的战略努力的初期，目的既是提供本土技术，也是打造一个新的出口部门。就军事开支而言，澳大利亚属于中等规模的参与者，2021—22年投入446.19亿澳元，占GDP的2.09%，占全球国防总开支的1.4%（相比之下，美国占39%，中国占13%）。但根据Morrison政府作出的承诺，这些数字将增长，其中包括在新的三边AUKUS安排下，为核动力潜艇投入至少700亿澳元。^[12]就出口而言，澳大利亚国防产业几乎还算不上参与者，2019—20年的出口额仅为55亿澳元。不过，这一数字已从2017—18年的15亿澳元增长而来，是一系列政府倡议、定向投资战略、学术研究资助计划和知识转移中心的直接结果。它们共同仿效深受新自由主义基础设施建设者喜爱的公私合作模式，旨在集中培育军事技术初创企业。为说明军事伦理如何在美德经济中发挥作用，本文认为，围绕“伦理”的三种关键动态正在塑造澳大利亚新兴军事技术产业及其与学术界的联系。首先，我以Cyborg Dynamics Engineering及其Athena AI平台为案例，考察伦理如何在军事技术初创企业中作为军事商品发挥作用。随后，我考察新的国防合作研究中心（DCRC）倡议的作用，聚焦位于昆士兰州布里斯班的“可信赖自主系统”DCRC，从而转向伦理在澳大利亚日益壮大的国防产业中所发挥的话语功能。第三，我有意提出一个具有挑衅性的命题：“伦理”促进了与大学的合作，而研究资助是其中的核心因素，并通过研究中心、网络、研讨会和工作坊具体化。在这一背景下，伦理成为国防—学术—工业复合体的关键词。最后，我主张，关于伦理和美德的军事观念，是美德经济中一个独特而又关键的组成部分，需要进一步研究和持续的批判性关注。

作为代码与商品的军事伦理

军事环境中充斥着伦理、法律和准则，从武装冲突法（LOAC）、行为准则，到战士精神等更为模糊却具有道德约束力的概念。国家自身也受到国际法的约束——至少在理论上如此，尽管实践中不总是如此。国际法既决定战争在何种情形下可以被认为是正义的，也保护冲突中平民的权利。当士兵奔赴战争时，“他们受到一系列伦理原则的约束，这些原则禁止某些行动，并严禁另一些行动。”^[13]有德性的行为未必写在这些原则之中，而可能是在特定军种或部队内部通过文化形成并维系的。我们可以将美德理解为与道德——良好品格——联系更紧密，而伦理则涉及行为，以及在特定情形下可允许行为的限度。军事伦理和美德都并非一成不变，而是随着时间发生了实质性演变，往往与新战争形式的兴起相伴，例如二十世纪空中力量的扩散。军事语境中的伦理与规范行为的准则和法律相混同，因此通常被作狭义理解；相比之下，士兵对于战士精神却可能感受到更复杂、更强烈的联系。

随着用于发现、选择、锁定和杀伤的技术出现，将伦理编码进计算系统的可能性，已被证明对军事领导者颇具吸引力。Christian Enemark观察到，无人机战争的兴起改变了国家对战争的伦理计算：士兵无需暴露于风险之中，平民伤亡据称也有所减少，这些都承诺通过技术进步产生更合乎伦理的战争。^[14]自主武器系统承诺，进一步从方程中剔除容易犯错且有缺陷的人类决策，将法律、伦理和行为准则转化为计算决策树；无论自主系统是否“智能”，都可审慎、灵活且可靠地应用这些决策树。在这里，伦理问题常取决于人类操作员与杀伤决策的关系——人在回路内、人在回路上监督，或人在回路外——或者更根本地取决于一种道德坚持：战争中的杀戮始终应由人类决定。^[15]如果自主武器系统的当前发展，以及2021年《特定常规武器公约》（CCW）未能对其使用实施监管可以作为迹象，那么将自动化杀戮视为道德上应受谴责的立场正面临严重困境。战争自主性的不断提高，无论就自主武器系统的实际运作而言，还是就支撑这些系统的算法过程与思维而言，都具有重大的哲学和实践意涵；它们有可能使法律遏制军事暴力的可能性本身不复存在，因为这种暴力先于并超越人类思考和决策的能力。^[16]

不过，虽然世界各地的学者、行动者与公众仍然忧虑“杀手机器人”的危险，机器是否会决定采取致命行动这个问题，在实践中大体上已成定局。在这里，军事伦理的独特性质，实际上促进了自主性不断增强的技术的出现。即使军事伦理被置于详尽的社会、制度、历史和哲学语境中（例如参见澳大利亚国防军2021年的《军事伦理》条令），它仍需要为了战场而被操作化，使士兵能够迅速作出生死决断。伦理一旦被理解为准则，便成为可编码之物——能够被翻译成计算形式，教授给智能系统，并应用于特定情境。正如Elke Schwartz所说，这种“伦理模块的逻辑，依赖于一种将伦理视为可以编纂、可以确定，并能产生清楚、安全、理想情况下确定无疑之结果的观念”。^[17]被编码的伦理成为商品，处于初创企业推销话术的核心，也对那些负责监督算法与自主技术开发和采购的军官具有极大吸引力。^[18]

澳大利亚初创企业Cyborg Dynamics Engineering提供了一个很有说明力的案例。它的旗舰产品是Athena AI，一个用于武器瞄准和战场分析的平台。公司网站称：

“Athena AI是市场上为数不多的基于视觉的AI系统之一，它把AI计算机视觉、由AI支持的决策辅助，以及在用户界面中显示AI信息结合在一起。”Athena AI本身虽然不是自主武器系统，但其设计目的是增强人的目标瞄准能力，关键还在于提供物体识别，以及伦理和法律评估工具。在公开演示和公司网站上，Cyborg Dynamics将这些伦理能力宣传为重要的差异化优势。Cyborg Dynamics创办人与任职或关联于“可信赖自主系统”国防合作研究中心的合作者，在一篇简短的反思性学术文章中，把该技术的目标描述为：“通过减少‘战争迷雾’、改善对国际人道法的遵守，增强人类在战场上作出伦理与法律决策的能力。”^[19]那篇文章展示了Athena AI开发过程中采取的迭代式、能够回应问题的伦理方法，而下面这段较长的公司网站引文，则揭示了伦理如何在营销修辞中成为附加价值：

Athena AI是为数不多的可信赖AI产品之一。我们曾与International Weapons Review、军事法律官员和军事伦理学者合作，帮助为AI视觉与决策辅助认证界定适当的数据保障和测试方法。我们的内置决策辅助工具，在适用情况下纳入了法律与交战规则方面的考量。

把武装冲突法、交战规则及其他此类准则与具体作战情形联系起来，以此实现伦理编码，被放在该系统优势列表的首位，使Athena AI被定位为相对于现状的一种改进。精品法律咨询公司International Weapons Review（IWR）的参与，也清楚呈现了美德经济中一些熟悉的组成部分。正如本书编者所描述的、卷入美德经济的技术批判机构，IWR这样的组织本身并不是“问题”——就IWR而言，该公司由有经验且具有扎实学术声誉的军事律师运营——但它们仍然是军事伦理经济多样且不断演化的版图的一部分。^[20]Cyborg Dynamics的修辞没有说出的是，平台各种检测和分类功能——据公司网站称，它们由多阶段神经网络实现——被该系统声称能够提供的法律架构隐含地合法化并放大了。从政治角度看，“伦理”在此通过将军事技术的关切转移至计算与法律相遇的技术官僚领域，降低了政治介入的强度。但在产业中，伦理发挥着提高系统商品价值的作用：系统本身具有美德，因为其中已经编码了伦理，因而它承诺使部署该系统的军队更加严格地遵守伦理准则。

伦理与研究成果转化的基础设施

Cyborg Dynamics这样的公司，是澳大利亚国防产业中不断壮大的中小企业

(SME) 生态的一部分。在前总理Malcolm Turnbull的保守派领导下，国防产业增长被视为至关重要的国家优先事项，既为了减少对外国进口的依赖，也为了创造就业。然而，澳大利亚国防军的采购规模不足以维持一个可持续的本土国防产业。正如澳大利亚国防部的《国防出口战略》所述：“需要新市场和多元化机会，才能帮助充分释放澳大利亚国防产业增长、创新及支持国防未来需求的潜力。”^[21]通过《2020年国防战略更新》，时任总理Scott Morrison承诺在随后十年投入2700亿澳元用于国防，旨在增强并更新澳大利亚军事能力，同时为产业与劳动者提供重大机会，并设计各种技能培训计划，支持海军造船及其他国防优先事项。

在这一推动之中，AI与其他高科技系统发挥重要作用，使澳大利亚国防产业被定位为具有创新性、面向未来，并已准备好为盟友的优先事项作出贡献。此方面的旗舰项目是波音“空中力量协作系统”（ATS）；这家美国国防巨头在其网站上将其称为“智能的无人化力量倍增器”。该项目与多家澳大利亚中小企业合作开发，昵称“忠诚僚机”，旨在研制一种高速攻击型无人机，能够为执行危险任务的人类飞行员提供支持，使飞行员与高强度冲突区域保持安全距离，或在空中缠斗时提供额外火力。因此，ATS必须在导航、制导和瞄准方面以相当高的自主性运作，而这又要求大量专业知识，并引出关于使用武力的重大伦理问题。

正是在此类情形中，“可信赖自主系统”国防合作研究中心（TAS）等研究成果转化机构发挥关键作用：它们既促进澳大利亚企业的参与，又在国防自主系统的设计与推广中凸显伦理。合作研究中心（CRC）是澳大利亚政府自1990年起实施的一项倡议，旨在将学术研究与“产业引领”的项目连接起来；资助通常以数千万澳元计，在数年间发放给至少包含一个产业伙伴和一个大学伙伴的合作体。一些CRC与国家安全有所交叉，例如“从数据到决策”（D2D）CRC和网络安全CRC，但它们更偏向民用关切；即使被纳入安全议程，也更可能涉及执法和信号情报等与国防相邻的活动。国防合作研究中心是较新的一类子项目，由“下一代技术基金”资助。根据工业、创新与科学部的一份情况说明，该基金在2016—17年至2025—26年间获得7.3亿澳元拨款，用于投资“与国防优先事项一致、面向未来且能够改变游戏规则的能力”。作为首个国防合作研究中心，TAS从澳大利亚政府和昆士兰州政府获得了大量资源与资金；昆士兰州也是唯一制定无人机产业战略的州。虽然波音ATS项目获得的媒体关注遥遥领先，但TAS的日常主要业务是与中小企业开展较小规模的项目，其中许多企业位于昆士兰州。不过，该CRC也在开展自己的倡议，为国防自主系统开发保障与伦理框架。

其中的核心组成部分，是与昆士兰大学“战争与法律的未来”研究小组共同开展的“可信赖自主系统的伦理与法律”项目。正如TAS网站所述，其“伦理提升计划邀

请多样化的利益相关者参与，为国防领域合乎伦理、可部署的AI提供以证据为基础且务实的风险管理”。与围绕大型科技企业的美德经济一样，该项目的目的并非质疑或批判国防产业的根本基础，而是“为部署在人机协作（HUM-T）中的电磁频谱、机器人、自主系统与人工智能的使用，制定伦理、法律、安全与问责框架”。项目由TAS首席科学家Kate Devitt博士领导；她受过伦理学训练，在数据与伦理方面拥有扎实的批判性学术研究经历。当然可以主张，应当积极理解此类倡议：它们在确保从一开始就把伦理纳入国防产业项目与产品方面，开展了至关重要的工作。按照这一解释，缺少这样的框架不会使倡议停止，只会意味着项目照常推进，而伦理在其构想中的位置更不重要。这些主张有其道理，而且TAS也利用其与澳大利亚国防军的密切关系，共同撰写了《国防领域伦理AI的一种方法》；这是国防科学与技术局（DSTG）的一份技术报告。DSTG是一个为新技术倡议提供资助、促进其发展并制作原型的军事机构。

在TAS的语境中，我们需要理解“伦理”至少以两种方式运作。第一种便是上文所述：TAS发挥基础设施作用，确保伦理考量成为澳大利亚所开发军事技术的基础。第二种则将这些伦理倡议视为在话语上使国防产业在学术界和更广泛公众中获得正当性的机制。在此意义上，“可信赖自主系统”国防合作研究中心可以被理解为某种伦理汇接中心，将法律、哲学及其他人文学科的研究与军事机构、从业者和产业连接起来，尤其关注初创企业。这样做使“伦理”能够被“内置”于AI和自主系统中，并在这些技术开发的同时或近旁，探讨军事情境下有关其伦理的棘手问题。TAS支持的学术出版物清单之广，足以证明TAS及其关联研究者认真对待伦理的理论和实践问题。^[22]TAS获得资助的目的也并非对军事行动或军事开支本身展开大规模批判，而是发展符合国防价值观与伦理的本土国防产业。正如创新环境中常见的情形，社会科学家和伦理学者的在场构成了一种照护工作；在这里，它变成对国家美德及其国防事业的照护。^[23]因此，TAS可以被理解为澳大利亚军事美德经济中的一个关键齿轮。这不是要一概否定最近一些倡议，例如部署AI识别西部Yalanji族人的文化遗产、协助保存约克角语言，或制定自主船舶实践准则，而是要认识到，所有这些努力都与美德的生产和商品化相互交织。

研究资助中的美德、学术界与“伦理”

过去二十年间，澳大利亚学术界的研究资助承受了越来越大的压力，2013年上台的保守派政府执政期间尤其如此。2014年，澳大利亚研究理事会（ARC）的竞争性研究资助为8.86亿澳元；八年后的2022年，这一数字为8.15亿澳元。作为有幸获得ARC资助的研究者，我能够证实，获得这类项目经费包含多少运气因素——在人文学科中尤其如此。政府越来越强调影响力、产业参与，尤其是研究商业化，使资金更加偏向应用研究而远离基础研究。2022年，Morrison政府宣布设立名为“澳大利亚经济加速器”的22亿澳元基金，其中16亿澳元指定用于可以“商业化”、并与“国家制造业优先事项”相一致的研究；这些优先事项包括“国防与太空”。在这一背景下，“伦理”成为描述人文与社会科学（HASS）研究角色的一个关键词，尤其在军事领域，因为它使研究者能够在原本集中于技术开发的研究空间中主张自身价值。国防与国家安全被表述为具有美德的事业，伦理也因此提供了一种共同语言，使人文社科可以与计算机科学、工程学、心理学，以及其他在历史上更紧密地与国防研究相连的学科交流。例如，人文学科关于传播的研究，可以帮助以正确方式进行信息战，或协助开发“合乎伦理”的自主战场系统。

要衡量军事资助对澳大利亚学术研究的全部影响，即使只作总体判断，也极为困难。在最基本的层面，就有如何对所谓“两用”研究进行分类的问题，例如美国国防部资助医学研究的情况。此外还有其他挑战。如何界定和划分国防、国家安全与情报资助？当没有明确预算可查时，能否识别由总理办公室、内阁乃至州总理资助的倡议？绝密倡议怎么办？预算金额与实际支出之间的差异又如何处理？虽然澳大利亚政府支出确实具有一定程度的透明度和问责要求，但由于国家安全部门的规模、保密性、复杂性与整体不透明，国防资助可能更难追踪。尽管如此，在澳大利亚，大部分较为公开、且相当可观的国防资助，来自国防科学与技术局

（DSTG）；这一国防部内部机构旨在协调研究优先事项，并为学术界和产业提供接口。“运筹研究网络”（ORNet）等合作倡议直接针对国防行动（指挥与控制、部队设计、作战规划等），而科学伙伴关系（DSP）计划则为与国防部门合作提供一个共同框架，国内每一所公立大学都列为其成员。在州层面，国防创新网络（新南威尔士州和澳大利亚首都领地）或国防科学研究所（维多利亚州）等组织，与DSTG紧密合作，围绕优先问题将大学研究者和中小企业连接起来。因此，学术界通过不断演化的制度基础设施与国防部门接触；这种基础设施用科学的语言包装国防优先事项，并提供论坛，使国防人员能够直接与多个领域的研究者交流。[译注：原书本段后一次将DSTG误作DTSG，此处统一为DSTG。]

美德与伦理很容易成为话语与情感层面的推动因素，使学术研究日益军事化。举例而言，颇具影响力的澳大利亚战略政策研究所（ASPI）公开倡导学术界与国防部门建立更加紧密的联系，并强烈强调国家安全的美德。虽然ASPI被设立为无党派智库，但其议程坚定地与不断扩张的国家安全体制保持一致；在澳大利亚境外，则

通过与美国、英国、加拿大和新西兰的“五眼”情报伙伴关系，与盟国保持一致。ASPI首席执行官Peter Jennings和前首席国防科学家Robert Clark在一系列博客、评论文章与报告中，呼吁建立“对五眼友好”的大学部门，并设立澳大利亚自己的DARPA。五角大楼著名的国防高级研究计划局DARPA，曾负责从早期互联网，到9/11之后把“地狱火”导弹改装到“捕食者”无人机上等各种创新。^[24]

大学往往急于承接新的国防倡议。我所在的机构在堪培拉运营澳大利亚国防军学院，并设有新南威尔士大学国防研究所；该研究所的网页充斥着粗粝的战争技术图像，实际信息却很少，甚至完全没有说明哪些人参与其中。不过，更多时候，国防倡议受到大力宣传，信息也很丰富。例如，我们的“可信赖自主性”研究小组和网络安全研究所，都是校内广泛宣传且十分活跃的实体；新南威尔士大学也在2018年凭借5000万澳元政府资助成立的网络安全CRC中占有重要席位。参与此类倡议的部分学者是我认识并尊重的同事；因此，我的意思再次不是指责个人。毕竟，即使我的研究没有国防资助，我所在大学也从澳大利亚国防军获得可观收入，并且像几乎所有澳大利亚大学一样，教学和研究都依赖国家资助。重要的是，这种实际运行的基础设施，对于应对当代国防挑战所提出的跨学科问题至关重要。在澳大利亚情境中，DSTG（国防科学、技术与研究局）的STaR shots计划最明确地表达了这些问题，包括“敏捷指挥与控制”“颠覆性武器效应”和“信息战”等主题。最后这一项，是研究自动化和AI的人文与社会科学学者尤其受到欢迎的领域之一，我收到的来自本院的参与申报邀请便证明了这一点。尽管我是一名批判军事技术及更一般军事化现象的人文学者，我仍常发现自己处在奇怪的场合：与获得国防资助的研究者，甚至国防人员一道参加工作坊、研讨会和讨论小组。对我而言，这些是观察系统内部、理解其动机与逻辑的宝贵机会。例如，我与“可信赖自主系统”CRC的接触让我获益良多，尤其因为它们让我深入看到，在军方周边及内部开展的细致、乃至具有批判性的工作。

然而，风险在于，在此类情境中，“伦理”变成人文与社会科学研究者能够为经费申请贡献的一项东西，而美德则作用于研究的框定方式：把此类研究说成可以拯救生命、保障繁荣安全未来的国家必需之事。在这些美德经济中，“伦理”含义游移不定，使它可以同时指计算机科学家开发的可编码规则，以及法律学者、哲学家和传播学研究者等所创造的过程、程序与论坛。因此，那些原本可能对从事“国防工作”感到不安的研究者，可以让“伦理”把自己与武力行动及致命暴力隔离开来，而后两者正是包括澳大利亚在内的世界各地军队的核心行动精神。这反过来服务于产业和国防部门的利益，因为它制造了美德的缓冲层，遮蔽眼前的残酷。“伦理”因而可以被理解为一种漂浮能指，一个具有可塑性的指称，它附着于许多情境，并且很容易被动员来服务军事目的。

结论

学术界、国防部门与产业之间的这种联系，本不应令人意外：大学一直都是军事性机构，服从军事目的，并浸透尚武政治。毕竟，相比作为学习、知识生产与异议的场所，大学首先更是帝国与殖民的机构。然而，这种军事性所采取的形式，会随时代、技术以及意识形态和经济观念而变化。本章论证，在军方、大学和产业围绕AI与自主技术汇聚的过程中，可以辨认出一种独特的经济形式：“伦理”根据情境作为商品或流通媒介发挥作用，“美德”则大量借用军方与国家主义的价值观。在此对一个以新军事技术为中心、正在演化的澳大利亚产业所作的勾勒中，“伦理”润滑了合作的齿轮，掩饰了国防的暴力目的，却又在实践中始终可以被化约为一组狭窄且可编码的规定；这些规定取自预先界定的有关武装冲突、武器和人权的法律与公约。当然，浅薄的“伦理”并不新鲜，但它在澳大利亚情境中所发挥的关键作用值得重视。描绘并分析AI研究、产业与应用的这一汇流，是一项关键任务，因为它所遵循的逻辑和资助经济，不同于科技产业支持AI与大数据学术研究时的通常模式。这一军事化的“伦理”构型有可能转移并蔓延至其他情境，尤其当资金匮乏的大学把目光投向澳大利亚公共生活中仅存的少数资金充裕机构之一时。

借用那个著名的问题：怎么办？鉴于硅谷与美国国防部之间技术转化的历史，大型科技企业日益嵌入美国军事体制并不令人意外，但这一过程并非一帆风顺。2018年的Google员工集体离岗抗议，部分由公司参与Project Maven引发，显示出一条潜在裂缝。高技能科技工作者比多数人更具流动性，其技能需求旺盛，因此相较多数行业的个人，他们拥有更强的议价能力。但尽管集体离岗抗议受到高度关注，Google也十分公开地在Project Maven问题上退让，科技与军国主义的根本关系并未发生实质变化。当昆士兰科技大学（QUT）宣布启动TAS倡议时，学生发起了#booksnottobombs（要书本不要炸弹）运动（图1），聚焦军事巨头BAE与Thales被纳入CRC资助模式的问题。虽然运动没有成功，但它与Google离岗抗议都表明，面对军方资金在学术界和科技界日益增强的影响，两界有必要作出集体回应。对在这一领域工作的批判性学者而言，一个关键步骤是追踪资金——不是去获取军事资助，而是描绘资金如何流经大学及其周边的学术体系。开展这种调查式工作，无疑需要集体劳动，以及建立新的知识网络。当然，实施这一项目不会逆转、甚至不会减缓这一特定美德经济的运作。但它可以揭示问题的规模，从本文呈现的一般性论断和具体实例的混合，走向更扎实的批判，说明美德如何在军方、学术界与AI研究的交汇处运作。



图1：“让昆士兰科技大学解除武装”（Disarm QUT）运动的横幅。图片来源：Monique Mann。

图内文字中文化：解除武装；面向真实世界的战争罪行；解除QUT武装；#要书本不要炸弹。原有被划去的导弹及Facebook标识保留。此图为原照片的AI中文化重绘，非历史现场的原始照片。

致谢

除为本文提供敏锐指导的编者与审稿人外，我还深深感谢Kate Devitt博士，她对早期版本提出了详尽、直率且建设性的反馈。Kate愿意回应一篇批评TAS的文章，证明了她自身严谨的伦理立场与宽厚的思想胸怀，但不应被视为她认可本文的论点。所有剩余错误均由我本人负责。

资助披露

本研究由澳大利亚政府倡议ARC DECRA DE190100486资助，也得到新南威尔士大学支持；该校是国防科学计划的成员，也是澳大利亚国防军学院的所属机构。

参考文献

Amoore, Louise. ‘Algorithmic War: Everyday Geographies of the War on Terror’, *Antipode* 41.1 (2009): 49–69. 中文题名：《算法战争：反恐战争的日常地理》。

- Amoore, Louise. *Cloud Ethics: Algorithms and the Attributes of Ourselves and Others*. Durham: Duke University Press, 2020. 中文题名：《云伦理：算法与自我及他者的属性》。
- Australian Defence Force. *ADF Philosophical Doctrine—Military Ethics*. Australian Defence Force, 2021, <https://theforge.defence.gov.au/adf-philosophical-doctrine-military-ethics>. 中文题名：《澳大利亚国防军哲学条令——军事伦理。澳大利亚国防军》。
- Australian Department of Defence, *Defence Export Strategy*, Commonwealth of Australia, 2018. 中文题名：《国防出口战略》。
- Beard, Matthew. ‘Beyond Tallinn: The Code of the Cyberwarrior?’ in Allhoff F, Henschke A, and Strawser BJ (eds) *Binary Bullets: The Ethics of Cyberwarfare*, New York: Oxford University Press, 2016, pp. 139–56. 中文题名：《超越塔林：网络战士的准则？》；《二进制子弹：网络战争的伦理》。
- Bellanova Rocco, Jacobsen, Katja Lindskov, and Monsees, Linda. ‘Taking the trouble: science, technology and security studies’, *Critical Studies on Security* 8.2 (2020): 87–100. 中文题名：《自找麻烦：科学、技术与安全研究》。
- Boeing. ‘Airpower Teaming System’, <https://www.boeing.com/defense/airpower-teaming-system/index.page>. 中文题名：《空中力量协作系统》。
- Buzan, Barry, Wæver, Ole and de Wilde Jaap. *Security: A New Framework for Analysis*. Boulder: Lynne Reiner Publishers, 1998. 中文题名：《安全：一种新的分析框架》。
- Clark, Robert and Jennings, Peter. ‘An Australian DARPA to Turbocharge Universities’ National Security Research: Securely Managed Defence-funded Research Partnerships in Five-Eyes Universities’, ASPI, 2021. 中文题名：《以澳大利亚版DARPA加速大学国家安全研究：五眼国家大学中受到安全管理的国防资助研究伙伴关系》。
- Department of Defence. ‘Memorandum for the Establishment of an Algorithmic Warfare Cross-Functional Team (Project Maven)’, 2017, https://www.govexec.com/media/gbc/docs/pdfs_edit/establishment_of_the_awcft_project_maven.pdf. 中文题名：《关于成立算法战跨职能小组（Project Maven）的备忘录》。
- Department of Industry. ‘Growth opportunities. Department of Industry, Science, Energy and Resources’, 2021, <https://www.industry.gov.au/data-and-publications/defence-national-manufacturing-priority-road-map/growth-opportunities>. 中文题名：《增长机会。工业、科学、能源与资源部》。
- Der Derian, James. *Virtuous War: Mapping the Military-Industrial-Media-Entertainment Network*, 第2版, New York: Routledge, 2009. 中文题名：《有德性的战争：绘制军事—工业—媒体—娱乐网络》。
- Devitt, S. Kate. ‘Normative Epistemology for Lethal Autonomous Weapons Systems’ in Galliot, Jai, MacIntosh, Duncan and Ohlin, Jens David (eds) *Lethal Autonomous Weapons: Re-Examining the Law and Ethics of Robotic Warfare*, Oxford: Oxford University Press, 2021, pp. 237–58. 中文题名：《致命自主武器系统的规范认识论》；《致命自主武器：重新审视机器人战争的法律与伦理》。

- Devitt, S. Kate, Scholz, Jason, Schless, Timo and Lewis, Larry (预印本). 'Developing a Trusted Human-AI Network for Humanitarian Benefit', *Journal of Digital War* “我的战争” (My War) 专刊, 2022. 中文题名: 《为人道主义利益开发可信赖的人类—AI网络》。
- Enemark, Christian. *Armed Drones and the Ethics of War: Military Virtue in a Post-Heroic Age*, New York: Routledge, 2014. 中文题名: 《武装无人机与战争伦理: 后英雄时代的军事美德》。
- Galliot, Jai, MacIntosh, Duncan and Ohlin, Jens David (eds). *Lethal Autonomous Weapons: Re-Examining the Law and Ethics of Robotic Warfare*, Oxford, New York: Oxford University Press, 2021. 中文题名: 《致命自主武器: 重新审视机器人战争的法律与伦理》。
- Halpern, Orit. *Beautiful Data: A History of Vision and Reason since 1945*, Durham: Duke University Press, 2015. 中文题名: 《美丽的数据: 1945年以来的视觉与理性史》。
- Howell, Alison. 'Forget "Militarization": Race, Disability and the "Martial Politics" of the Police and of the University', *International Feminist Journal of Politics* 20.2 (2018): 117-36. 中文题名: 《忘掉“军事化”: 种族、残障及警察与大学的“尚武政治”》。
- Liljefors, Max, Gregor Noll, and Daniel Steuer (eds). *War and Algorithm*, London; New York: Rowman & Littlefield International, 2019. 中文题名: 《战争与算法》。
- Massingham, Eve. 'Automation of the Spectrum, Automation and the Spectrum: Legal Challenges When Optimising Spectrum Use for Military Operations', *Law, Technology and Humans* 3.2 (2021): 91-106. 中文题名: 《频谱的自动化、自动化与频谱: 优化军事行动频谱使用时的法律挑战》。
- Ministers for the Department of Industry, Science, Energy and Resources. 'Action Plan to supercharge research commercialisation', 2022, <https://www.minister.industry.gov.au/ministers/taylor/media-releases/action-plan-supercharge-research-commercialisation>. 中文题名: 《加速研究商业化行动计划》。
- Packer, Jeremy and Joshua Reeves. *Killer Apps: War, Media, Machine*, Durham: Duke University Press, 2020. 中文题名: 《杀手级应用: 战争、媒体、机器》。
- Phan, Thao, Jake Goldenfein, Monique Mann, and Declan Kuch. 'Economies of Virtue: The Circulation of 'Ethics' in Big Tech', *Science as Culture* 31.1 (2021): 1-15. 中文题名: 《美德经济: “伦理”在大型科技企业中的流通》。
- Roberson, Tara, Stephen Bornstein, Rain Liivoja, Simon Ng, Jason Scholz, and Kate Devitt. 'A Method for Ethical AI in Defence: A Case Study on Developing Trustworthy Autonomous Systems', *Journal of Responsible Technology* 11 (2022). 中文题名: 《国防领域伦理AI的一种方法: 可信赖自主系统开发案例研究》。
- Sadowski, Jathan. 'Potemkin AI.' *Real Life*, 6 August 2018, <https://reallifemag.com/potemkin-ai/>. 中文题名: 《波将金式AI》。
- Schwarz, Elke. *Death Machines: The Ethics of Violent Technologies*, Manchester: Manchester University Press, 2018. 中文题名: 《死亡机器: 暴力技术的伦理》。
- . 'Silicon Valley Goes to War: Artificial Intelligence, Weapons Systems, and Moral Agency', *Philosophy Today* 65.3, (2021): 549-69. 中文题名: 《硅谷参战: 人工智能、武器系统与道德能动性》。

- Sharkey, Noel. 'Automating Warfare: Lessons Learned from the Drones,' *Journal of Law, Information and Science* 21.2 (2011): 140–54. 中文题名：《战争自动化：从无人机获得的教训》。
- Shepherd, Tory. 'Australia's Aukus Nuclear Submarines Could Cost as Much as \$171bn, Report Finds' *Guardian*, 13 December 2021, 世界新闻版, <https://www.theguardian.com/world/2021/dec/14/australias-aukus-nuclear-submarines-estimated-to-cost-at-least-70bn>. 中文题名：《报告发现，澳大利亚AUKUS核潜艇成本可能高达1710亿美元》。
- Suchman, Lucy. 'Algorithmic Warfare and the Reinvention of Accuracy', *Critical Studies on Security* 8.2 (2020): 1–13. 中文题名：《算法战争与精确性的重新发明》。
- Suchman, Lucy, Karolina Follis, and Jutta Weber. 'Tracking and Targeting: Sociotechnologies of (In)Security', *Science, Technology, & Human Values* 42.6 (2017): 983–1002. 中文题名：《追踪与瞄准：（不）安全的社会技术》。
- Viseu, Ana. 'Caring for Nanotechnology? Being an Integrated Social Scientist', *Social Studies of Science* 45.5 (2015): 642–64. 中文题名：《照护纳米技术？成为融入项目的社会科学家》。
- Whittaker, Meredith. 'The Steep Cost of Capture', *Interactions* 28.6 (2021): 50–5. 中文题名：《俘获的高昂代价》。

原注

- [1] Thao Phan et al., 'Economies of Virtue: The Circulation of 'Ethics' in Big Tech,' *Science as Culture* 31.1 (2021): 121–35. 中文题名：《美德经济：“伦理”在大型科技企业中的流通》。返回
- [2] Department of Defence, 'Memorandum for the Establishment of an Algorithmic Warfare Cross-Functional Team (Project Maven),' 26 April 2017, https://www.govexec.com/media/gbc/docs/pdfs_edit/establishment_of_the_awcft_project_maven.pdf. 中文题名：《关于成立算法战跨职能小组（Project Maven）的备忘录》。返回
- [3] James Der Derian, *Virtuous War: Mapping the Military-Industrial-Media-Entertainment Network*, 第2版, New York: Routledge, 2009, p. xxxi. 中文题名：《有德性的战争：绘制军事—工业—媒体—娱乐网络》。返回
- [4] Der Derian, *Virtuous War*, p. 83. 中文题名：《有德性的战争》。返回
- [5] Orit Halpern, *Beautiful Data: A History of Vision and Reason since 1945*, Durham: Duke University Press, 2015. 中文题名：《美丽的数据：1945年以来的视觉与理性史》。返回
- [6] Jathan Sadowski, 'Potemkin AI,' *Real Life*, 6 August 2018, <https://reallifemag.com/potemkin-ai/>. 中文题名：《波将金AI》。返回
- [7] Meredith Whittaker, 'The Steep Cost of Capture,' *Interactions* 28.6 (2021): 50–55. 中文题名：《俘获的高昂代价》。返回
- [8] Phan et al., 'Economies of Virtue,' p. 2. 中文题名：《美德经济》。返回
- [9] 例如参见：Louise Amoore, 'Algorithmic War: Everyday Geographies of the War on Terror,' *Antipode* 41, no. 1 (2009): 49–69; Louise Amoore, *Cloud Ethics: Algorithms*

and the Attributes of Ourselves and Others Durham: Duke University Press, 2020; Rocco Bellanova, Katja Lindskov Jacobsen, and Linda Monsees, 'Taking the Trouble: Science, Technology and Security Studies', *Critical Studies on Security* 8.2 (2020): 87–100; Jeremy Packer and Joshua Reeves, *Killer Apps: War, Media, Machine*, Durham: Duke University Press, 2020; Lucy Suchman, 'Algorithmic Warfare and the Reinvention of Accuracy', *Critical Studies on Security* 8.2 (2020): 1–13; Lucy Suchman, Karolina Follis, and Jutta Weber, 'Tracking and Targeting: Sociotechnologies of (In)Security', *Science, Technology, & Human Values* 42.6 (2017): 983–1002. 中文题名: 《算法战争: 反恐战争的日常地理》; 《云伦理: 算法与自我及他者的属性》; 《自找麻烦: 科学、技术与安全研究》; 《杀手级应用: 战争、媒体、机器》; 《算法战争与精确性的重新发明》; 《追踪与瞄准: (不)安全的社会技术》。返回

- [10] Alison Howell, 'Forget "Militarization": Race, Disability and the "Martial Politics" of the Police and of the University', *International Feminist Journal of Politics* 20.2 (2018): 117–36. 中文题名: 《忘掉“军事化”: 种族、残障及警察与大学的“尚武政治”》。返回
- [11] Barry Buzan, Ole Wæver, and Jaap de Wilde, *Security: A New Framework for Analysis*, Boulder: Lynne Reiner Publishers, 1998. 中文题名: 《安全: 一种新的分析框架》。返回
- [12] Science Department of Industry, 'Growth opportunities,' Text, Department of Industry, Science, Energy and Resources, Department of Industry, Science, Energy and Resources, 30 March 2021, <https://www.industry.gov.au/data-and-publications/defence-national-manufacturing-priority-road-map/growth-opportunities>; Ministers for the Department of Industry, Science, Energy and Resources, 'Action Plan to Supercharge Research Commercialisation,' 2 February 2 2022, <https://www.minister.industry.gov.au/ministers/taylor/media-releases/action-plan-supercharge-research-commercialisation>; Tory Shepherd, 'Australia's Aukus Nuclear Submarines Could Cost as Much as \$171bn, Report Finds,' *Guardian*, 13 December 2021, 世界新闻版, <https://www.theguardian.com/world/2021/dec/14/australias-aukus-nuclear-submarines-estimated-to-cost-at-least-70bn>. 中文题名: 《增长机会》; 《加速研究商业化行动计划》; 《报告发现, 澳大利亚AUKUS核潜艇成本可能高达1710亿澳元》。[译注: 原书日期作“2 February 2 2022”, 重复的数字2照录。] 返回
- [13] Matthew Beard, 'Beyond Tallinn: The Code of the Cyberwarrior?' in *Binary Bullets: The Ethics of Cyberwarfare*, Fritz Allhoff, Adam Henschke, and Bradley Jay Strawser (eds), New York: Oxford University Press, 2016, p. 139. 中文题名: 《超越塔林: 网络战士的准则?》; 《二进制子弹: 网络战争的伦理》。返回
- [14] Christian Enemark, *Armed Drones and the Ethics of War: Military Virtue in a Post-Heroic Age*, London: Routledge, 2014. 中文题名: 《武装无人机与战争伦理: 后英雄时代的军事美德》。返回
- [15] Elke Schwarz, *Death Machines: The Ethics of Violent Technologies*, Manchester: Manchester University Press, 2018; Noel Sharkey, 'Automating Warfare: Lessons Learned from the Drones', *Journal of Law, Information and Science* 21.2 (2011): 140–54. 中文题名: 《死亡机器: 暴力技术的伦理》; 《战争自动化: 从无人机获得的教训》。返回

- [16] 例如参见：Max Liljefors, Gregor Noll, and Daniel Steuer (eds), *War and Algorithm*, (London; New York: Rowman & Littlefield International, 2019. 中文题名：《战争与算法》。返回
- [17] Schwartz, *Death Machines*, p. 16. 中文题名：《死亡机器》。〔译注：原书此处及正文将姓氏拼作Schwartz；本章完整书目作Schwarz。〕返回
- [18] Elke Schwarz, ‘Silicon Valley Goes to War: Artificial Intelligence, Weapons Systems, and Moral Agency’, *Philosophy Today* 65.3 (2021): 549–69. 中文题名：《硅谷参战：人工智能、武器系统与道德能动性》。返回
- [19] Tara Roberson et al., ‘A Method for Ethical AI in Defence: A Case Study on Developing Trustworthy Autonomous Systems’, *Journal of Responsible Technology* 11 (2022): 1. 中文题名：《国防领域伦理AI的一种方法：可信赖自主系统开发案例研究》。返回
- [20] Phan et al., ‘Economies of Virtue’, p. 10. 中文题名：《美德经济》。返回
- [21] Australian Department of Defence, *Defence Export Strategy*, Commonwealth of Australia, 2018. 中文题名：《国防出口战略》。返回
- [22] 例如参见：S. Kate Devitt, ‘Normative Epistemology for Lethal Autonomous Weapons Systems’ in Galliot, Jai, MacIntosh, Duncan and Ohlin, Jens David (eds) *Lethal Autonomous Weapons: Re-Examining the Law and Ethics of Robotic Warfare*, Oxford: Oxford University Press, 2021, pp. 237–58; S. Kate Devitt et al., ‘Developing a Trusted Human-AI Network for Humanitarian Benefit,’ (预印本) ; Jai Galliot, Duncan MacIntosh, and Jens David Ohlin (eds), *Lethal Autonomous Weapons: Re-Examining the Law and Ethics of Robotic Warfare*, “伦理、国家安全与法治”丛书 (Ethics, National Security, and the Rule of Law), Oxford; New York: Oxford University Press, 2021; Eve Massingham, ‘Automation of the Spectrum, Automation and the Spectrum: Legal Challenges When Optimising Spectrum Use for Military Operations’, *Law, Technology and Humans* 3. 2 (2021): 91–106; Tara Roberson et al., ‘A Method for Ethical AI in Defence: A Case Study on Developing Trustworthy Autonomous Systems’. 中文题名：《致命自主武器系统的规范认识论》；《致命自主武器：重新审视机器人战争的法律与伦理》；《为人道主义利益开发可信赖的人类—AI网络》；《致命自主武器：重新审视机器人战争的法律与伦理》；《频谱的自动化、自动化与频谱：优化军事行动频谱使用时的法律挑战》；《国防领域伦理AI的一种方法：可信赖自主系统开发案例研究》。返回
- [23] Ana Viseu, ‘Caring for Nanotechnology? Being an Integrated Social Scientist,’ *Social Studies of Science* 45.5 (2015): 642–64. 中文题名：《照护纳米技术？成为融入项目的社会科学家》。返回
- [24] Robert Clark and Peter Jennings, ‘An Australian DARPA to Turbocharge Universities’ National Security Research: Securely Managed Defence-Funded Research Partnerships in Five-Eyes Universities,’ *Australian Strategic Policy Institute blog*, 14 July 2021, <https://www.aspi.org.au/report/australian-darpa-turbocharge-universities-national-security-research-securely-managed>. 中文题名：《以澳大利亚版DARPA加速大学国家安全研究：五眼国家大学中受到安全管理的国防资助研究伙伴关系》。返回

第三部分：行动

“公开的秘密”：Meredith Whittaker 访谈

访谈：Jathan Sadowski、Thao Phan

Jathan Sadowski (JS)：先从你的经历谈起吧。你是怎么进入这个领域，从在谷歌工作，到共同创办 AI Now 的？^[1]

Meredith Whittaker (MW)：我算是误打误撞进来的。对我而言，谷歌就是一份工作，不是什么使命。我想，这个视角帮助我形成了立足现实的看法：从劳动的角度看科技问题，把学术界看成一个行业。我在伯克利读书，2006年本科毕业。当时我身无分文，谷歌则是一家正在招人的本地公司。那时谷歌正在迅速扩张。“向右上方增长”^[2]仿佛会永远持续下去；你会听到 Sheryl Sandberg 这样的人一本正经地说，谷歌已经“破解”了资本主义永远增长、同时又造福世界的秘密。这让人反感，却又十分引人入胜。而且很难向我的朋友解释。这肯定不符合我在“食物而非炸弹”^[3]的同志们的理念。但与此同时，眼前看不到任何反向叙事。“科技已经找到答案”的感觉占据了霸权地位。我就这样被扔进那个环境，作为合同工为 Writely 做客户支持；Writely 是谷歌收购的一项产品，后来成了 Google Docs。

刚开始，我年薪4.5万美元，比我父母任何时候挣得都多。这很令人兴奋，但它并不是我的身份，我也没有下意识地努力去爱它。我原本就没有热爱工作的倾向。我把雇佣劳动看作一种骗局。我出身的阶级不会把工作当成自己的身份。工作是你讨厌、却又不得不做的事，做得越少越好；而生活，则是工作结束后发生的一切。

JS：2006年的谷歌还完全是一家正在崛起的巨头。我可以想象，刚毕业不久就有那样的收入、那样的声望，很多人一定会很高兴地说：“迷魂汤在哪里？我要一饮而尽。”

MW：是啊，那是个疯狂的地方。他们提供免费果汁，每间洗手间都有那种每个大约1000美元的 TOTO 智能洁身马桶。显然，这套东西对某些人很奏效。这也是当时如此引人入胜的原因之一。我只是试着把一切弄明白，问一些天真的问题，比如：到底是什么赚了这么多钱？这些钱从哪里来？整理信息……这究竟是什么意思？

有一段时间，我能在谷歌不断上升、发展顺利，原因之一恰恰是我的出身完全没让我为那种环境做好准备。我不懂专业精英那套礼貌政治。我没有意识到，谷歌的空白支票其实并不是开给我的。我追逐“20%时间”^[4]之类的福利，还相信那些关于扁平层级、任人唯贤、无论出自何处最好的想法总会胜出的企业叙事。这让身边的

人陷入尴尬：要么把我叫到一旁，解释职场中隐含的不平等如何运作，要么就让我过关。因此，我毫无心机地骑着谷歌的 Gbikes 在园区转悠，去其他楼里见人，提出20%项目，总之是睁大眼睛、满怀期待地试着兑现那张空白支票。2006年的谷歌有许多自由空间，而我不知道自己不该占用它们，于是我就占用了。

这还真奏效了！我最终被调到工程部门的一支新团队，负责文档标准，并开始带领一个小组。那份工作其实冲突十足，让我得以近距离观察企业怎样通过标准机构和志愿组织，以有利于自己的方式塑造技术标准，同时维持集体专家共识的表象。正在崛起的科技行业寄生于这些较早的组织形式，并将其收编，而这一过程在外部观察者眼中并不明显。此后，我又参与创办和运营 M-Lab、谷歌开放研究小组、AI Now 等组织。

Thao Phan (TP)：你能说说自己总共在谷歌工作了多少年吗？

MW：我从2006年7月开始，最后一天是2019年7月。所以，是13年。

TP：你觉得是否有某个时刻真正改变了你对当时状况的看法，或者发生过什么特定事件？你一直都感到疏离，还是在某个时刻会说：“其实现在已经不只是疏离了。现在这是不道德的。它违背了我个人的伦理”？

MW：有好几个转折点。我拿的是英语文学与修辞学学位，当初选择它，是因为伯克利那些酷孩子都在那里。我并不理解在那里接触到的所有东西，但有一种情感上的吸引：那些怪人在这儿，而我爱他们。我会读这些看不懂的玩意儿，因为我是在一群看上去善良、古怪又很酷的人组成的共同体里读它们。那是21世纪初，后结构主义流淌在学术界的血液里，大家仍然认真对待精神分析；我几乎没有地图，就被扔了进去，读着拉康，怀疑自己到底还会不会阅读。我困惑极了，简直像个孩子。

所以，我接触过很多从结构上思考权力的方法，但当时这些路径在我的日常生活中并没有具体的参照物。我想，后来正是这些思想，成了我理解自己在谷歌所经历之事的工具。就在那儿，我开始重新阅读福柯、马克思主义理论，以及早期的无政府主义思想家——大学里没有教过他们，但我原本就有所了解。随着我开始辨认权力的运作方式，这些思想突然在我心中亮了起来。

要形成一套让我自己信服的、关于谷歌和科技的分析，是个漫长而充满冲突的过程，而且它仍在继续。不过，我想自己原有的批判倾向，以及没有经历传统计算机科学学位训练这一点，都对我有所帮助。我的性别不对，阶级出身不对，也不熟悉企业职场中那些用来标示归属的仪式。因此，我不断与疏离感搏斗，也必然会对那些被别人视为公认常识的事情提出疑问。这在我 and 那碗迷魂汤之间筑起了一层防护。

刚入职时，我连服务器是什么都不知道。我对计算技术一无所知，而且，就像对公司整体作出的承诺那样，很多东西我都照单全收。比如谷歌“整理了全世界的信息”这一说法。好吧，真酷，大家都在点头，那一定就是谷歌在做的事情了！但为了理解那些我以为别人都懂的事，我会问一个非常愚蠢的问题，比如：什么是信息？而后，往往会露出一个巨大的窟窿，因为我逐渐消化了这样一个事实：所谓信息，指的其实是采用 HTML 格式、能够被网络爬虫访问的内容。我会围绕这些问题和发现反复思索，它们就像手里不停拨弄的念珠。但我没有可以一起这样思考的知识共同体，因此这个过程既孤独，又断断续续；为了应对它，我读了很多书。当我得知存在一个叫“科学技术研究”的领域时，真的几乎哭了。我找人交谈，已经找了太久！

JS：你刚才谈到这只是一份工作时，说到了一点，我觉得非常好，而且其力量比乍看之下还大。就是说，你上班打卡，下班打卡，你他妈的讨厌自己的工作；而生活是工作之前和之后发生的事，不是在工作的那段时间里。但我觉得，现在相信这一点，其实是一种低调的激进立场，尤其在我们这个领域，尤其是在马克斯·韦伯所谓的“志业”之中——那些被视为“天职”的职业。我认为，科技行业和学术界都把自己抬高到这种地位：这些不是工作，而是使命；你必须热爱自己所做的事，必须全身心投入。

然而，另一面却从没有人说出来：“是啊，可它并不回报你的爱，也不回报你的投入。”我想，把它视为志业而非工作的观念，也阻止了许多人成为自己职业内部的批评者，或提出那类问题。你能否进一步谈谈，把它认作一份工作究竟激进在哪里，以及为什么我们从事的职业往往缺乏这种意识？

MW：这个问题我想得很多，因为显然，我在谷歌的这段人生历程最终落在了对劳工组织工作的投入，以及对这些问题的追问上：谁是科技劳动者？什么是科技工作？我们他妈的究竟在做什么？而且，我想，这个问题表明，我们需要进一步追问：这些机构如何能够劫持或塑造我们的自我感、归属感，以及共同体和交往的空间，再把这些引导成一种要求——要求我们表现出几乎像家人一样的忠诚。

如果人们不把一个光鲜的专业头衔看成使命或身份，会有什么利害关系？如果像我对待洗碗或零售工作那样，把它当成一份工作，他们会失去什么？究竟会失去什么？学术界和科技行业都是如此。我现在身处学术界，与学者交谈时发现，他们基于职业的自我评价，与科技行业极其相似。

对我而言，组织工作如此激进、又确实令人快乐的地方，在于拆解我们与雇主、与工作的感情联系，再把这些感情和忠诚转向身边的人。这就是建立团结的过程。当你更多地通过自己所爱的人和彼此的纽带，而不是自己在机构等级中的位置来认同自己时，团结才存在。在共同体和团结中建立的身份，可以取代通过职业晋升阶梯和绩效评估建立的身份。这不会一夜之间发生。它是一种反思性实践，要求我们拆

解自己的自我感，以及它对我们在这些机构中地位的依赖，努力通过彼此的关系，以及把你和同事作为人连接起来的友谊与爱的纽带，重新界定自己。

我认为，这个过程必须从承认这些都是工作开始。只要他们付钱给你，归根结底就可以告诉你该做什么，无论这种权力关系被多少层障眼法包裹。事实上，没有人会忘记是谁给自己发工资，永远不会。在工作中获得某种奖励——无论是美国国家科学基金会（NSF）的资助、晋升，还是终身教职——并不意味着你是更好的人。它意味着，你让那些掌握着支配你的权力的人满意了。这样说，就从根本上拆穿了那种靠拿高分过活的生活。

JS：我真的听到过太多学者郑重其事地说，他们工作的资金或支持来源并不影响他们的工作，简直不可思议。就是说，他们确实把自己看成安·兰德式的超人，完全不受任何外部影响，超脱于权力斗争之上，只做自己的事。别人资助他们，只是因为喜欢他们本来就在做的研究。

MW：对，而我们又可以问：如果他们不再这样相信，会面临什么利害关系？就我个人而言，我这一辈子没有一秒钟忘记过是谁给我发工资，因为如果没有工资，我就活不下去。这是不会忘记的。担任 AI Now 的学术主任时，我也感到过冲突。比如，我正在写一篇东西，而我知道它会刺中某根敏感神经，会激怒很多人，会把事情揭开到一种程度，让那些大学在资金上所依赖的科技企业非常痛恨。同时，我必须顾及自己对 AI Now 博士后、对那些需要从 AI Now 走向就业市场的人的照护责任。如果我或者 AI Now 让某些不该得罪的捐赠人或大学董事感到不舒服，研究所的资金、声望，以及支持其成员的能力，会受到什么影响？

TP：作为劳工组织者，当你发现自己处于那样的权力位置，能够为别人创造工作、给别人就业保障时，一定是个很有意思的时刻。这是一种巨大的责任。但你觉得，它会不会变成一种借口？变成说服自己接受道德上有问题的资金的理由，比如军事经费或来自采矿业的产业资金。我知道，很多学者觉得自己的主要责任是为别人创造工作。在这个充满不稳定、伦理上又处处妥协的资助环境里，我们能做的最好事情，就是创造工作、重新分配资金。

MW：换一种说法，你就是个老板。在大学的语境里，在我的处境中，你还是一名中层管理者；你的工作是筹来资金，以不会让捐赠者不安的方式提高学校的声望，从而让大学行政部门满意。而在这些大学里，你实际上并没有那么多权力。上面是一层又一层收入丰厚的行政管理者：教务长、副教务长、院长、副院长。这里没有像产业中那样直接的老板，也没有以透明方式强制执行劳动关系的结构，因此对劳动者的规训是间接发生的，在很大程度上通过内化的自我规训来实施；对那些受过训练、善于考试、习惯拿高分的人来说，这似乎更容易。他们非常擅长预测怎样才能讨评分者的欢心。从很肤浅的层面看，这或许像是独立——例如，我写作时并没有人真地站在我肩膀后面盯着——但实际情形远非如此。而且，我认为，个人对于

独立和能动性的感受，即认为学者出于对知识的纯粹献身而工作、不受基本生存忧虑干扰的假定，削弱了我们认识这些结构对自己施加多么强烈影响的能力，最终使组织起来和发展劳动意识变得极其困难。

JS：学术界存在一种奇怪的心理学和社会学现象。我一直说，太多学者都把成为中层管理者当作志向。他们的目标是在一个更大的组织内部，掌握对一小块领地的一点权力。我想，这也体现在大学中层管理职位的大量增加上。这还极大地遮蔽了其中运作的政治经济关系：没有人是劳动者，每个人都是管理者。

MW：而这一切建立在不稳定就业经济的背上，建立在承担大学大部分实际教学工作的兼职教师队伍之上。这就是今天的大学。少量相对保全下来的岗位仍然存在：有保障、有医疗保险、有些声望的终身教职轨道岗位。而获得这些越来越稀少的岗位的机会，有助于强迫成千上万的人顺从；他们竞争的，可能真就只是本领域里的两三个职位。如果那一年找不到工作呢？真遗憾，你就加入兼职教师的后备队伍吧。这就是为什么我认为，若不以劳工组织工作为中心，就不可能解决困扰大学的问题，因为眼前发生的是赤裸裸的权力较量。他们不会因为我们提出了一条更好的定理，就给予我们更多权力、能动性和保障。

JS：我确实想谈谈 AI Now，以及你从谷歌内部的批评者转变为整个行业外部批评者的过程。有个从军事领域借来的词，叫“红队演练”。大型组织很喜欢设红队。他们就像友好的批评者，在战争、市场或其他场景中扮演敌对一方，让我们更好地理解对方的策略，以及如何反制。但说到底，内部批评者或红队仍然必须是这个集体的一员，仍然必须忠于组织。

MW：我很喜欢红队这个比喻。我看到这些公司采用的正是一种遏制策略：给我这样的人一个平台，却不给我们影响决策的权力。在我看来，微软研究院也发挥着类似的作用。这样，公司就可以声称自己咨询了批评者、综合了我们的意见，从而给它原本就打算做的事涂上一层合法性的薄膜。我得承认，我花了一段时间才意识到这一点。

我在离开谷歌很久以前就共同创办了 AI Now。因此，从某种意义上说，我本人就是这些冲突的一个具体例子。不过，我也会认为，自己在某些方面颠覆了它们，或者至少试探了它们的边界：我把谷歌的大笔资金引向公司以外的批判性工作，再把其中一部分转给其他从事激进工作的人。但是，作为一名仍受雇于科技公司的劳动者，竟然能够直接创办一所大学研究机构，这种能力值得在这里仔细考察；它表明大学与科技行业之间在结构上相互渗透的程度有多深。大学，特别是计算机科学和工程学科，依赖科技企业的资金和基础设施。这也意味着，提出 AI Now 的想法时，没有人觉得有什么奇怪。

要理解这一切是怎样发生的、怎样走到今天这个地步，我们可以看看“AI”和机器学习的近代史。只要回溯15年，“AI”领域还很偏僻冷清。只有少数坚守者研究机

器学习及相关学科。他们会在圣迭戈之类地方的中档酒店里，参加小型会议。他们没有自己的实验室，也没有运营什么光鲜的研究中心，主要是因为这些领域当时并不被认为特别重要或能赚钱。随后，大约在2010年前后，科技经济中的赢家开始巩固地位；这些公司凭借监控广告商业模式致富，并在科技行业中占据越来越核心的位置。

像 Facebook、谷歌这样的公司，原有的组织结构就让它们拥有处理和收集数据的庞大基础设施，拥有巨大的市场覆盖面，因此可以获取不断更新的数据流。这种覆盖面和获取能力，不是随便打个响指、花钱就能买来的。所以，它们手里有了这些要素：集中化的基础设施、数据，以及持续为它们创造新数据的市场渠道。而那个重要的顿悟时刻在于，人们发现，把那些较早的机器学习技术置于如此大量的已标注数据和如此庞大的计算基础设施之中，它们就能做到以前做不到的事情，也就是在某个基准测试中取得更好的经验性表现。但“AI”技术本身并不新；新的，是数据、基础设施和覆盖能力的集中。

这让科技企业猛然醒悟。它们意识到：嘿，我们有数据，有算力；现在，凭借已有几十年历史的机器学习技术，我们又有了一种可以浇在这些东西上的神奇酱汁，可以用它创造“智能”产品和服务，为我们进军广泛的市场提供理由。

现在，这些公司能够声称，它们正在构建通用、智慧、智能、超越人类等等技术；这些技术会在无数领域增强或替代决策，同时当然也进一步巩固它们的权力。

那么，这与学术界和科技行业之间广泛的纠缠有什么关系？因为搞“AI”的人本来没多少，大多数人都待在某所大学里落满灰尘的实验室中，安安静静做自己的事。到了2010年代初中期，一场招聘热潮突然出现，这些人收到的薪资报价堪比足球运动员。各家公司突然都在争抢“AI”研究者，因为它们意识到“AI”非常赚钱。而争抢“AI”研究者，也意味着争相资助“AI”研究；这种研究需要强大的计算基础设施，需要获取大量标注数据，才能以当时被认为前沿的方式加以利用。因此，双重任职制度开始流行：研究者一边保留大学的职位与头衔，一边为科技企业工作。关键还在于，转向“AI”意味着，大学若想在计算机科学和 STEM 学科排名中保持竞争力，就必须争抢数据和基础设施，而这些东西掌握在科技行业手中。我在《俘获的高昂代价》^[5]中更详细地讨论了这种被俘获的关系，梳理了它的结构，并分析为什么这种纠缠会严重危及学术独立和批判性工作。

JS：你能谈谈 Maven 计划^[6]、谷歌员工罢工，以及那段时间发生的其他事情吗？

MW：Maven 是谷歌与五角大楼签订的一份秘密合同，旨在为无人机的目标锁定与监视开发机器视觉“AI”。那是一份小合同，但其设计意图是充当入口，帮助谷歌做好准备，在竞标 JEDI 合同时获得更大机会。JEDI 是一个庞大的项目，要为军方提供从战场到指挥站的全部计算基础设施和“AI”。

谷歌很有策略。它请来了 VMware 的前首席执行官 Diane Greene，后者又从自己的政府关系团队中带来了大批人员。我在2017年底得知 Maven 合同的存在；当时我的工作，是说“当心 AI，它既虚假又古怪”，然后希望有人愿意听。

接着，我看着 Maven，心想：我被耍了。我们全被耍了。这他妈的太恶劣了。既然我靠着所谓的批判挣工资、赢得声名，怎么可能不为这件事把自己押上去？我是说，我们谈的是谷歌：一家仿佛掌握了我们最敏感信息档案的公司，知道我们的朋友、购物、位置。现在谷歌还想用天知道什么烂数据训练机器视觉系统，给美国非法的无人机战争配备更多自动化监视和目标锁定系统？谷歌是一家海外员工比美国国内员工还多的跨国公司，现在却想为美军工作，建立一种依赖关系，让美国政府越来越有能力向谷歌施压，利用这些联系？不行。

就在这时，我和另外一小群人转向了劳工组织工作。显然，我们不可能靠观念说服他们。他们知道那些观念，听过我和其他人提出严密、精心组织的论证。我们是对的。这不是问题所在。因此，我们需要建立一股真正能够对抗其权力的力量。我完全不知道这是否可能。那时，对我们少数几个人而言，这只是一次试验。但我们尝试了。我写了一封公开信，即“取消 Maven”的信，获得了数千人签名。那是2018年初，它触动了人们的神经。因为至少在当时，公司里还有许多人真心相信自己在服务公共利益，并认为谷歌不该像信中所说的那样，从事“战争生意”。公司没有预料到这一切；至少在最初大约三个月里，我们一直远远走在它前面，占据了优势。

我们只是尽可能把事情向前推进，即便谷歌领导层一直试图介入和遏制，有时反而弄巧成拙。比如，他们举办了一场向全球播出的员工大会，把我飞送到山景城，让我上台参加三场辩论：第一场早上8点开始，最后一场晚上8点开始，以覆盖全球各地。令我震惊的是，没有一位高管做了准备。我只提前48小时收到通知，但仍写了好几页笔记，做过演练，把所有数字和统计资料都整理在提示卡上。而他们走上台时，竟然毫无准备。因此，他们的傲慢和本可避免的失误，在许多方面帮了我们。幸好，那时我凭借自己的工作和声誉，已经积累了足够资本，使别人愿意认真对待我。

我们成功取消了那份合同，而且我想，也点燃了人们组织起来的愿望；随后，他们开始参与，为建立更持久的力量打下基础。Maven 事件结束时，我知道，想继续留在谷歌，自己已经完蛋了。我知道他们会等一阵子，不会一下子全面出手。但他们会来收拾我，因为我不再只是一个唱反调的声音。我真的让他们付出了代价，而他们已经不觉得这有什么可爱了。于是我想，自己唯一能做的，就是把油门踩到底，继续组织，看看会发生什么，同时详尽地做记录，一刻也不放松警惕。

Maven 是第一波公开可见的、针对当今科技行业形态的精英科技劳动者组织行动。它向科技行业的神经系统注入了一种观念：组织起来，是制约这些强大公司、

实现某些“伦理”结果的正当且有用的手段；否则，这些结果只会躺在架子上，成为从未落实的立场文件。员工罢工，是这种新近获得认可的抵抗形式最醒目的表现。那实在太太好了。行动很快成形，主要由呈现女性气质的劳动者以及运营岗位上的人组织；他们的日常工作本来就是整天把各种事情组织起来，可以说正是维系公司运转的黏合剂。突然之间，数百名极其能干、主要呈现女性气质的人，把每天靠以领工资的工作技能转过来对付公司，用另一种方式运用所有这些才干。

大约三天，一个完整的组织系统就建立起来了：各地行动团队向全球每个地点派发写有要点的传单，地区负责人向核心组织者汇报，分发网络确保每个人都得到最新的行动规程，还有一组人不断修订诉求、收集反馈。从零开始，就建立了一整套成熟的基础设施。光是谈起它，我都觉得受到鼓舞。这真正展示了集体行动的可能性，也展示了如果把劳动转向别处，我们能够做成多少事情。它还改变了人。人们很快乐！我们突然清楚地看到了两个世界的差别：在一个世界里，我们把精力和社会关系投入为谷歌赚钱；在另一个世界里，我们把它们投入创造我们想要的世界。我看到人们开始问：“我们还能做什么？”这些让重新夺回自身劳动的可能性显露出来的时刻，力量极其强大。谷歌恰恰在罢工之后^[7]才开始认真镇压，绝非偶然。

TP：我想，现在能否把话题转向机构中的硬权力与软权力？你在产业中工作时看到了这些问题，后来转到学术界，却又看到同样的权力通过资助、通过其他影响形式发挥作用。你能带我们看看，大型科技企业怎样对学者施加影响吗？

MW：我在谷歌时，资助关系友好的学者是常规做法。政策团队经常会与友好的研究者一起塑造一份白皮书或学术论文，而这些研究者也很可能正在接受你的资助。随后，当这篇友好论文发表，其中包含你想要的图表，表明YouTube对音乐人有益，或者别的什么能让你赚更多钱的结论时——也就是说，当它作为事实进入了权威知识体系——你会得到表扬，因为你可以把这个小事实放进Eric Schmidt在欧盟委员会面前发表的讲话之类的场合。这些事情并不隐秘。它们是常规做法，是政策、公关传播和游说策略的关键部分。

我现在是一名学者，但大学对我的待遇，从来没有比我还是一个手握资金的科技劳动者时更好过。而当我开始在纽约大学全职工作时，我看到了另一面：大学中的政治经济如何运作，项目拨款和资助对于成功有多么重要，又如何取决于你是否从事某种不那么激进的工作。在谷歌时，我知道公司拉拢学者是为了影响他们，但我并没有意识到这种影响究竟具有多大的权力。这些公司的控制简直令人窒息，尤其是在新自由主义大学寻找大捐赠者、普遍不愿惹恼这些捐赠者的情况下。

我和AI Now都因此受到重创。纽约大学工程学院——我们当时设在那里——警告我们，要让自己的观点与其他工程研究中心保持一致；而那些中心总体上是实证主义和技术官僚取向的。2021年初，我们被告知，自己不符合学院的战略，必须另找归宿。随后，2021年底，纽约大学通知我们，要拿走我们所有的捐赠资金。留

给我们的选择是：要么打官司，可诉讼很贵，而作为纽约大学一部分的 AI Now 又无法支付诉讼费用；要么离开，再想办法补上实际上被大学夺走的四百多万美元。

我们仍在处理这件事的后果，未来还不确定。同样，我们也不可能靠取得更多成功来摆脱这种处境。这一切发生时，按任何衡量标准，AI Now 都是世界上最有影响力的、从事批判性科技研究的大学研究中心之一。我们与少数其他人一起塑造了一个领域。我们不仅达到了所有指标，而且在大学表面上看重的那些方面，以多个数量级超出了指标。媒体报道全都有我们，等等等等。但到了关键时刻，这些并不是驱动新自由主义大学的东西。说到底，我所能得出的唯一解释是：工程学院知道，如果得不到亚马逊赞助的机器学习中心，得不到 Meta 的基础设施拨款之类的资源，就不可能提升排名。这些是开展“前沿”计算机科学研究的基本条件，而 AI Now 挡了路。我想，对批判性工作的威胁，已经再清楚不过了。对于这些大学，我们能够争取到、用来雇两名博士后的福特基金会25万美元资助，与 Eric Schmidt 刚刚给耶鲁大学某个新 AI 中心的1500万美元，根本无法相比。它们很清楚这一点。我们没有多少筹码。

Amy Westervelt 有一个很好的播客，叫 *Drilled*，讨论气候否认主义的历史。她考察了这样一段时期：埃克森曾资助、推动严谨的气候研究，希望或许能改变商业模式，转而供应非碳能源。反正两条路都能赚钱，何乐而不为？但到了某个时刻，它们决定，自己其实并无兴趣探索那些非碳能源。大约就在那时，它们决定把气候科学家赶出去，最终转而资助气候否认论，为其提供平台，并资助能够支撑这种说法的边缘研究。这改变了气候研究领域，制造出一种环境：公开发声的批评者成为打击目标，面临职业生涯受威胁的处境。我看到，科技行业也处在一个类似的转折点。有一阵子，它们以为自己吸收批评并把它重新纳入体系。那个时刻已经过去，而批判性工作现在处境十分岌岌可危。鉴于此，我们需要更多空间来讨论和制定策略；我担心，资助者召集的那些彬彬有礼的会议室，并没有提供这样的空间。

TP：我想回到劳动的问题。资助的政治经济，怎样影响了今天机构内部劳动的性质？很多学者现在觉得，为了吸引资金和其他机会，自己必须像品牌一样运作。这不是冒犯，但你也算是这个领域中的一位微名人。不过我想，即便你刚才举的 AI Now 的例子——超出了所有关键绩效指标，却仍受所属机构威胁——也说明，同行中的声誉还不够。你还必须拥有合适类型的品牌和伙伴关系。

MW：对。我当然是一个公众人物，但这件事给我的感觉，像是它自己就发生了。把自己变成品牌这整套东西，既棘手又令人不快。它让结识新人、与他们建立关系更加困难，因为作为一个有感情的普通人的你，与别人从推特和只言片语中认识或想象的那个你，是不同的。无论如何，我想，这是无休止、令人精疲力竭的注意力与认可竞争的一部分。我们希望，这种竞争哪怕能换来零星资助也好，因为我们依靠这些钱创造一些稳定性，让工作得以继续。成为微名人是一种策略。撤走社会工

资、实行不稳定劳动制度，就会出现这一切；而如今，多数学术工作都属于这种制度。

我看到，它催生了一种不断积累成就和赞誉的文化，试图用越来越多的成功压过不稳定和不安全的力量。这很有害，与“模范少数族裔神话”和任人唯贤的神话有许多共同之处：它把成就和成功看作可靠指标，试图不断积累它们，以追求某种身份和地位；而这些身份和地位与阶级、白人至上主义的关系，远比与“才能”的关系密切。这种冲动——为了获得医疗保险而门门拿 A——是组织工作的敌人。组织工作需要劳动者意识，需要我们作为人对自己、对彼此诚实，说出：这他妈的一层层垒高的成就文化，对谁都没有好处。它把我们变成竞争者，而不是同志；哦，对了，它也没能保护我们的安全。有权力的人并不是微名人，他们不为人所知，却享有隐私和安全的好处。他们不必走到台前跳舞，只为换取某个基金会丢来的一枚 25 美分硬币，并承诺完成一大堆交付成果。我还应该指出，右翼并不采用这种模式。科赫家族、默瑟家族，以及从科技行业流入的资金，数额可观，也有保障；它们不要求人们每年为争取经费而跳舞，而是给予丰厚资助，让他们做自己的事，而这些事恰好有利于当权者。

这整套安排还会制造一种循环的倦怠：你永远不够安稳，无法扎下主根，坐下来开展那些安静、批判性强、需要长时间酝酿的工作。相反，我们都在追逐最新的流行语，指望把圆棒硬塞进方孔，好让自己的人继续有工作。客气地说，这产生的是与时代需要不相称的工作。过去大约五年，我们就看到所有人都跳上“AI”的潮流，而且这样做也情有可原，因为那里还有一丁点钱，或许能让人文社会科学继续有片遮风挡雨的屋顶。

TP：很高兴你说出了这些。我想，听到 AI Now 在自己所属的机构里处境艰难，很多人都会非常震惊。在澳大利亚尤其如此，因为正如你说的，现在 AI 太热门，伦理太热门，而这也是人文学科里为数不多还能找到工作的地方。Jathan 和我，都受雇于某种“AI 促进社会福祉”类型的中心。

我们正在接受这样一个事实：如果不以这种方式定位自己，就没有我们的工作。我想，许多人文学科机构和学者，都在一个极其敌意的环境里，拼命奔向 AI 伦理，把它当作生存手段。所以，即便我们中最成功的 AI Now 也在艰难挣扎，这一点一定会让许多人震惊。

MW：这令人震惊，情感上也非常难熬。我想，一旦异议所针对的对象弄清楚怎样管理这些异议，那些原本强有力的异议者逐渐退出、或被排挤出某场运动、某个空间，其实并不少见。就 AI Now 而言，我们是非常有力的批评者。我把在科技行业内部十多年间知道的一切，都带进了这项工作；这些知识，在学术界并不容易获得。尤其是过去一年左右，我们毫不留情。我们直接批评 Eric Schmidt，而且直指资金利益。我知道他们为什么不喜欢我们，因为我们确实有效地扭转了其中一些

叙事，使那些掌握巨额资本、巨大权力的人感到不便。纽约大学不会与我和几个基金会站在一起去对抗 Schmidt；顺便说一句，那几个基金会始终非常支持我们。但大学可以通过把我赶走、让这些工作噤声，来向 Schmidt 示好。

JS：我觉得，这次访谈中讨论的大部分内容，可以看作学术界和产业界一系列公开的秘密。在这些空间工作的人，能够看到俘获和影响正在发生，也看到它们对我们的工作、对我们所在的组织产生的直接和间接后果。但是，即便承认越来越多批判声音正在揭露这些动态，更常见的做法仍然是什么也不说。我想，这有好几个原因。也许我们正在从这些关系中获益，或者希望有一天能获益；也许我们不觉得现行安排有什么不妥，或者想象不出事情还能怎样；也许我们觉得，说出来在职业上显得无礼，在个人层面又令人不舒服；又或者，正如你刚才谈到的，Meredith，我们害怕公开发声会给职业生涯带来后果，而且我认为这种担心有充分理由。

考虑到这一切，我们该怎样做必须做的事情，也就是你在工作中实际正在做的事——面对俘获、收编与妥协的动态？我们怎样才能既正面又立即地开展这项工作？

MW：有个听起来有点轻率的回答：在一定程度上，做就是了。你把话说出来，就是说出来，然后去洗手间吐一场。你就是去做。但我想，还有……我觉得自己有某种义务，要把这些事情说清楚，因为凭借难以置信的运气和偶然的环境，我恰好既曾身处谷歌内部，从一侧看到这一切，后来又进入了那种被科技资金层层包裹的精英学术界。我处在一个非常有利于理解这些资本流动如何运作的位置，而且坐在最前排。

所以，我觉得有责任把自己知道的东西重新投入知识的公共积累之中；这些是我因特殊位置而掌握的信息，从学术界内部大多看不到。如果我不这样做，那就不叫团结，也不是真正利用我所知道的、以及这些经历为我的分析带来的东西，去推动任何一个我愿意参与的未来。如果我们不说清俘获的程度，不认真承担绘出其结构的任务，我想，我们就处在非常危险的位置。沉默并不能保障安全——这座金字塔骗局正在崩塌。除非改变它，否则，登录学术零工平台、用比同事更低的报价争抢一份兼职教学任务，就是我们的未来。对太多人来说，这已经是现在。承认这些动态，是组织工作能够开始的地方，也是我们可以开始把同志关系置于职业人脉之上的地方。

当然，这很难，也会让人变得脆弱。对于那些喜欢拿 A 的成年人，对于那些把安全感和自我价值建立在自己能够不断获得权势者嘉奖的能力之上的人，尤其如此。

我们还必须正视其中的利害关系。雇佣劳动支配着生活的方方面面。人会因为没有工资而死去，他们所爱的人也会死去。这不是夸张。有些人无法冒险站出来，也不该被要求如此——例如患有慢性病、需要医疗保险的人，持签证、工作资格可能被剥夺的人，等等。我们必须辨认自己需要哪些团结关系，以及学术界的斗争如何与其他斗争相互联系。有力的社会工资，会真正帮助解决其中许多问题。如果我们有

稳定的工作，就不会忙着微调那些他妈的会议论文里的资助披露段落；我们根本不会以现在这种方式，如饥似渴地追逐经费。因此，我们的运动、我们的组织工作，有一部分需要把注意力放在这些根本问题上，确保人们能够更安全地组织和冒险。我们想要一个这样的世界：人们不必在说出真实情况与付得起胰岛素的钱之间作选择。

原注

- [1] 本次访谈进行于Meredith担任美国联邦贸易委员会AI高级顾问、继而出任Signal总裁之前。
[返回](#)
- [2] “向右上方增长”是硅谷用来描述强劲增长曲线的一种说法。[返回](#)
- [3] “食物而非炸弹”（Food Not Bombs）是一场反资本主义的全球社会运动，抗议人为制造的浪费和匮乏。该运动特别关注食物浪费，收集被丢弃的食品，将它们做成餐食提供给社区。
[返回](#)
- [4] 这是谷歌倡导的一项员工福利制度，保证员工可以将20%的时间用于自主安排的额外项目。
[返回](#)
- [5] Meredith Whittaker, 〈俘获的高昂代价〉（‘The Steep Cost of Capture’），*ACM Interactions* 28. 6 (2021): 51–55。[返回](#)
- [6] 关于 Maven 计划的进一步讨论，参见本书 Richardson 的文章。[返回](#)
- [7] 谷歌“为真正的改变而罢工”（Walkout for Real Change）是谷歌员工组织的一系列同步离岗抗议活动，针对有关性不当行为、性别不平等和系统性种族主义的指控发声。示威发生于2018年11月，全球有超过20,000名员工参与。[返回](#)

“退出者”：Lilly Irani、Alex Hanna、Khadijah Abdurahman 访谈

访谈：Jake Goldenfein

本次访谈的参与者因共同决定退出一场会议而建立联系：这场题为“科技领域的权力与问责”的会议，由加州大学洛杉矶分校（UCLA）技术、法律与政策研究所于2021年11月主办。在这里，他们反思“退出”的政治，并将其与本书的核心主题相联系。请记住，《美德经济》的分析工作，是理清研究者在与产业资本互动时面临的复杂困境和悖论；而退出，则回应了产业资金政治的另一个维度。许多学者通过直接招聘、资金支持或演讲平台寻求与产业互动，退出却展现了一种承诺：要剥去这样一种收编行为的合法性，即把学术论坛变成产业利益积累知识资本与声誉资本的工具。产业支持的活动往往放大与自身契合的观念，但批判性学者的参与，仍然为其提供知识上的正当性。批判声音可能给有问题的观念披上学术外衣，帮助构建一种分析公正、参与者多元、论证严谨的框架。因此，批判性学者的劳动成了产业利益借助学术权威、推动支持自身议程的政策工具。当然，论坛的政治性质至关重要，产业资金也并非全都一样。那么，什么时候适合参与，什么时候适合公开批评，什么时候又适合退出？

Jake Goldenfein (JG)：先请大家自我介绍吧。

Lilly Irani (LI)：我叫 Lilly Irani，是加州大学圣迭戈分校的副教授。这几年，我一直与亚马逊 Mechanical Turk 平台的劳动者一起组织行动，努力改善他们的工作条件。更早以前，很长一段时间里，我和 Six Silberman 共同为该平台劳动者运营一套软件系统。我还参与组织科技劳动者联盟。在进入学术界之前，二十多岁时，我在谷歌那座“让世界更美好”的教堂里工作。当时我没有真正意识到，那里是一个数字基础设施帝国的触角。如果那时已经有科技劳动者联盟，也许我会留下来开展组织工作；但在二十多岁时，组织工作完全不在我的自我理解和社会知识之中。

Alex Hanna (AH)：我叫 Alex Hanna，是谷歌伦理 AI 团队的一名研究科学家。

^[1]我主要研究数据和机器学习如何加剧种族、性别与阶级不平等。目前，我还没有一个明确的组织工作归属，但过去做过劳工组织工作、争取跨性别医疗服务的组织工作，也参与过“科技伦理”领域中围绕具体倡议展开的行动。

J. Khadijah Abdurahman (KA)：我是 Khadijah Abdurahman。从职业上说，我是谁呢？我会说，自己研究儿童福利系统中的预测分析；坦白说，这个问题大多数时候根本没人关心。我觉得，科技领域的人压根不考虑儿童福利或对家庭的警务

管控。我与学者一起做了很多组织工作，较少涉及劳工权利。在某些方面，我试图改变话语，甚至试图改变人们对“什么是话语”的理解。

JG: UCLA 会议召开前，Khadijah 联系了我，指出这所具体的 UCLA 研究机构从查尔斯·科赫基金会获得了高达400万美元的资助。网络检索显示，其他慈善资助方，如 Brad Jones、Troesh 家族基金会，以及 Anthony 和 Jeanne Pritzker 家族基金会，也作出了贡献，使研究所的总捐赠资金大约达到1000万美元。（奇怪的是，查尔斯·科赫基金会的资助文件如今已无法在网上找到；会议举行时，它肯定还在网上，但最近几次搜索都找不到了。）

Khadijah 联系我时，我有些不知道该怎么办。我是一个在世界另一端工作的、相对默默无闻的学者。科赫基金会的政治，并不是我天然就理解的事情。所以，我担心自己的退出很可能毫无意义。我想，与其退出，是否不如利用分配给我的小组发言时间，谈谈科赫和产业资助。但当 Lilly 和 Alex 都在这件事上率先行动后，退出就更有意义了：它是向你们二位表达团结的一种姿态。

能否请各位简要说说，自己最初怎样成为活动参与者，最后又为什么退出？

LI: 我收到参加邀请的方式，与过去十年里似乎无穷无尽的那些工作坊一样：总是在谈怎样用某种数字东西来修复某种东西。UCLA 研究所的主任联系我，请我谈科技劳动者行动主义。我说：“你们不是应该请 Meredith Whittaker 吗？因为她真的曾在谷歌组织反对 Maven 计划的行动。”^[2]那件事里，我作为学者发起联署，以团结支持的方式参与。他说：“哦，是的，我已经请过 Meredith 了，但她实在参加不了，所以才来请你。”我说：“哦，那很好，我们想法一致。”唉，我错得多离谱。接受邀请后，Khadijah 联系我，指出活动接受了科赫的资金。我后来与 Khadijah 和 Meredith 交谈，Meredith 说：“哦，我其实知道，所以我才拒绝参加。”这一点，我此前并不知道。

对我而言，退出几乎是无需犹豫的决定。原因也不只是我的工作有保障，还因为我已经在 Meredith 曾称作“批判性 AI 产业复合体”的环境里待了十年：各种活动得到资助，聚在一起谈伦理问题；大批基金会、政府组织和大学都想借此打响名号、获得经费，也许还能影响话语。由于我与亚马逊 Mechanical Turk 劳动者一起工作，过去十年，我参加了一场又一场讨论数字劳动问题的类似会议。最后，我得出的结论是：这些论坛根本他妈的改变不了什么。在某些方面，与这些公司的研究者、律师、政策人员建立人脉，其实是在消耗精力；这些精力本该用来建立团结，与受到最直接影响、对这些系统该如何改变也最有发言权的人一起组织。

所以，第一，这些事情会耗费精力；第二，它们会开始把那些希望改变平台、却还没有行动计划的人圈在其中。于是我想，与其花三个小时参加这场小组讨论，又为该说什么而焦虑，不如把这些时间用来和 Turkers——Mechanical Turk 平台劳动

者——一起组织，向人们解释科赫资助的问题，或加入并放大 Khadijah 这样的声音。

生命有限，一天只有那么多小时。我已经决定，更严格地选择把自己的时间交给谁，为哪种团结关系投入时间。

JG：感谢你给我们这一小时，我很珍惜。

AH：我也是以差不多的方式参与这场工作坊的。我收到了邀请，此前从没听说过这个中心。我问了 UCLA 的另一个人，这个中心是做什么的；对方似乎愿意替它背书。我并不想具体点出那个人是谁。然后，我也觉得：“好吧，不妨参加。”活动在推特上公布后，Khadijah 同样联系了我，提醒我注意科赫资金。其实我本来就有些怀疑这场讨论，因为它的框架是怎样在企业内部追究企业责任，而相关参与者来自 BSR [企业社会责任组织，Business for Social Responsibility] 以及企业社会责任领域。由于它的局限相当明显，那并不是我真正想参加的对话。

Khadijah 提醒我之后，我便决定：“好，对，我退出。”于是我离开了这场活动。我完全同意 Lilly：这些事情有自己的经济体系，其中存在一种“声望经济”，与围绕成为那个吹哨人、成为那个让科技企业承担责任的人的讨论紧密相连。但站在高台上做这些事，有着相当明确的局限。如果不理解这种动态，却把精力投入其中，就会挤压那些基于亲身遭受伤害而拥有更深刻分析之人的空间——那些正在组织的人，那些承受不成比例的监视与监禁性控制的人。

KA：我没有被邀请在这次活动上发言。我也在想，自己经常不被邀请参加各种活动。我不一定认为这是因为我在学术等级中的位置，而是因为我常常会说：“你们他妈的在干什么？”刚才我其实在笑，因为 Timnit Gebru 被解雇、或被“辞职”、不管该用什么词之后，^[3]Alex 和我本来还要参加另一场小组讨论，主题同样是科技劳动者组织。最初，他们只告诉我 Alex 会参加，我就想：“好啊，当然，我愿意。”后来我才知道，另一位将参加讨论的人，不仅不是科技劳动者组织者，还是个名望很高、狂热的白人至上主义者。于是我说：“绝对不行。我不会参与，因为这完全是一场骗局。”由于活动在公立大学举办，他们不能公开撤回对别人的邀请，只好取消整场活动。我觉得这没什么问题，因为我为什么要为这种事出卖自己？对我而言，我并不认为存在某种你能抵达的位置，让你在道德上完全不受牵连。但只要我还有能动性，为什么要毫无理由地与一个漫画式恶棍同台？我的这种作风确实出了名，而且相当广为人知，所以大概也导致很多事情不会邀请我。

至于 UCLA 的活动，我此前关注过 Kate Klonick 在《纽约客》上发表的一篇关于 Facebook 监督委员会的文章。我记得，由于她长期嵌入 Facebook 内部，她受到了猛烈批评。那大约也是当时总统的妻子 Jill Biden 高度强调别人应称她为“拜登博士”的时候。人们作出类比，说 Kate Klonick 作为女性遭到不公平针对，也就是说，这是厌女，她被要求遵守不同标准。而我想：不，我相当确定，那是因为她

向 Facebook 出卖了自己，正在兜售企业版本的政权宣传。后来，我看到 David Golumbia 的一些推文，才弄清楚 Kate 接受了查尔斯·科赫基金会技术与创新部门的领导职位。对我而言，这令人震惊。因为，的确，我们所有人都被卷入了这些没有道德、充满肮脏金钱的世界；但科赫是非常具体的一种情况。科赫在围绕批判种族理论的讨论、向环境排放毒物等方面都极为系统，而且还创造了一套让学术研究被抽空一切意义的流程。他们系统性地运用资金，创造一个解除监管的政策空间，让资本能够进一步形成垄断。^[4] Klonick 接受科赫的钱，这一点对我尤其刺眼。

接着，我看到 Kate Klonick 被安排在 UCLA 法律与政策研究所关于平台问责的会议上发言，发现活动和研究所都由科赫资助，于是联系了你们。我还要说——我不会点出具体姓名——因为我们都在这种彼此接近的关系网络中活动，连我也与 UCLA 有关联，所以我认识一些说“不”的人：他们拒绝退出。我们 [Khadijah 和 Jake] 还讨论过，退出是否仅仅是在标榜美德。对我而言，我觉得这种想法很危险。我不认为，每一个公开作出的道德决定都是表演。我确实认为，我们在世界上的所作所为是有意义的。一方面，如果有人就在此刻去微软参加活动，我或许不会抱有同样的感受。但我认为，有些时刻，我们拥有更大的能动性。比如人们因 Maven 计划而在谷歌罢工，或者围绕 Timnit 发生的事情。有时候，会出现一个让局面发生巨变的契机。

AH：对于 UCLA 的活动，我当时在想：“对，我们有很多朋友也会做这样的事。”他们确实是我们视为盟友或朋友的人。这里存在一种张力；当人们从某些地方接受资金时，你必须认真思考，这个联盟到底是什么样子。我觉得科赫是一个特别尖锐的案例，但我也想到谷歌解雇 Timnit 后我们的那些讨论：我们开始请求人们停止接受谷歌的钱。有些人照做了，Queer in AI 做了，Black in AI 做了，我记得 Widening NLP (WiNLP) 也做了。但也有人回来对我说：“好吧，我们从谷歌得到的就这么一点钱。拒绝它，会让我们的活动项目处于不利地位。”

我想继续追问这种张力。这里没有一条绝对界线，但重要的是问：这种张力会产生什么作用？我想讨论拒绝或退出的行动，以及它们在特定时刻能做什么。我认为，这里有必要展开分析，因为对我而言，并不总能清楚判断，什么时候撤回参与或退出的效果最强。我很想听听你们的想法。

KA：我在想，资金怎样组织共同体。我们还没有发布这次访谈，但我曾就运营知识生产中心时的转化式正义问题，采访过 Safiya Noble 和 Meredith Whittaker。^[5] Meredith 的一句话让我一直记得：“什么是共同体？”这个词如今与 Facebook 的联系很强；但很多时候，所谓共同体——“批判科技”这顶大伞——其实只是阿斯彭研究所或某些资助者召集到一起的那群人。群体中的某个人，可能不只是接受科赫资金，实际上还支持科赫、自由意志主义及相关的一整套东西；另一个人，则可能以被监禁的黑人女性为关切中心。许多以共同体、友谊或亲近关系

名义流通的东西，都是十分人为地生产出来的。疫情期间，这一点更加令人迷惑，因为对普通人而言，以活动、出版物或某种“话语”呈现的这些关系，其实际条款和条件并不真正可见。尤其是，话语和学科往往由机构创造，而这些机构未必向我们负责。

JG：在我们 [Jake 和 Khadijah] 的交谈中，你提出了一个很有说服力的观点：在科技政策/AI 伦理领域，科赫基金会的资金尚未成为常态。它还不普遍，也不必成为不可避免的事。

KA：在这个具体案例中，我觉得围绕科赫资金和学术问责，存在一个让局势发生巨变的机会，尤其是因为科赫在这个特定领域的霸权还没有被巩固。而且，活动内容本身也有些可疑：难道我们通过一场小组讨论，就能让平台承担责任？

因此，在这个《爱丽丝梦游仙境》般的兔子洞里，我确实觉得，这是我们可以推动改变的地方；它也确实让人不舒服，尤其是因为我与一些没有退出的人是朋友，而他们没退出让我很生气。他们大概也烦了：我算谁，是个因为别人参加邪恶讨论会就羞辱他们的学术保姆警察吗？但我确实认为，我们必须对自己会做什么、不会做什么，划出一些明确界线。这很难，因为权力运作的许多方式都被隐藏了，或者由于我们共同的共谋而变得模糊不清。再加上我们正处于疫情之中，大家都他妈的压力巨大：哪里该妥协，哪里该推动，哪里又该要求别人也站出来？

LI：有一种说法是，资助者太多了，情况令人难以承受，根本不可能处理。但并不是说，一旦画下一条明确界线，它就永远是那条界线。因为对我个人而言，那条界线是资本主义。坦白说，问题在于，拒绝科赫是否让我离终结资本主义更近了一步。但每一次划线，都会迫使人们讨论：为什么界线在这里，未来又可能需要哪些界线。这不只是做出退出的姿态，也是创造条件，让更多人能够增进知识、深化分析、组织起来，并认识到我们在撤回劳动或开展其他直接行动时，实际拥有怎样的力量。

我同意 Khadijah 的看法：在她刚才精彩阐述的那些方面，科赫的确很独特。我也觉得，有必要把这条界线带到其他地方。但如果我们连科赫这一点都不能达成一致，就不可能走到那些更微妙、更友善、研究也更不充分的地方，去画下未来必须画出的明确界线。

JG：能否再谈谈特定的“时刻”或具体的明确界线？与其他那些给我们钱的人相比，科赫真的恶劣到如此不同吗？

KA：这里就是张力！我很喜欢你们那篇《美德经济》论文，^[6]部分原因是它提出了 Minderoo 的问题。我很清楚自己接受了 Minderoo 的资助，虽然是通过 AI Now、以多少有些间接的方式。我以前没听过 Minderoo，因为我不是在澳大利亚长大的。但这东西真的邪恶透顶。它建立在剥夺原住民土地的基础上，Twiggy

Forrest 是第六代澳大利亚殖民者，还尤其支持通过算法分配福利给付，这与我的研究直接相关。在美国人眼中，Minderoo 并不像我想象中澳大利亚人所看到的那样，是一个显而易见的邪恶实体。我知道自己正在接受有问题的钱，但我也处在资源极度匮乏的环境中。我们这里像是在经历大萧条；我不只是说我个人，而是说，我与一些人一起工作，始终在想办法给他们转去资金。我不愿要求他们无偿做事，也觉得有责任替他们找到钱。对我而言，这就是张力的一部分。

谈到划出明确界线，对我来说，有一条围绕着警察。我可以坐在这里与你交谈，Alex，也可以批评谷歌；但如果你为 ICE [美国移民与海关执法局] 工作，我们就不会出现在这通 Zoom 电话里。我相信反过来也一样。这又更加复杂，因为我们知道谷歌与国防部纠缠在一起。我特别提起这一点，是因为 Lilly 和我曾参与给 FAccT^[7] 中央委员会写信——“中央委员会”听起来太马克思列宁主义了，反正不管它叫什么——请他们公开表明对警察的立场：“你们到底支持不支持警察和警务研究？”他们拒绝画下这条明确界线，而这对我来说是非常清楚的一件事。

JG：我记得2019年隐私法学者会议（PLSC）及其与 Palantir 的关系。记得当时有一封公开信，要求会议取消 Palantir 的赞助资格；组织者回答说：“嗯，你知道……我们有赞助政策，规定赞助者不得施加任何影响，所以我们还是更愿意让他们参与讨论。”最后，学术界形成了足够一致的压力，迫使会议放弃了这家赞助者；但这发生在尽管有那份赞助政策的情况下，而不是由于那份政策。

我认为，这种回应代表了这些学术空间去政治化的更广泛问题。一切都变得相当程序化。你会在会议赞助政策里看到类似说法：“只要赞助者不影响活动或研究所的内容与议程，一切就没问题。”这样一来，我们画出不可跨越的明确界线的能力，甚至思考这些界线的能力，就被抽空了；因为已经没有空间，让你采取伦理立场，或者根据自己在这方面的伦理态度，发展作为研究者的主体性。

LI：Jake，你提到大学去政治化时，我脑子里一直浮现学术自由的问题，因为我所在的大学现在也在发生相关的事情。

科赫兄弟为自己在大学中所做的事情辩护时，有一种说法就是称之为“学术自由”。而如今，学术自由已经变成了保护种族主义、厌女和法西斯言论的空间。我记得读过，20世纪三四十年代的美国，学术自由是大学给予教师和研究生的东西；这些人基本上就是在社会主义旗帜下组织工会的大学劳动者。后来，我认为，美国大学教授协会（AAUP）接受了一种薄弱版本的学术自由与终身教职，以此取代对大学更实质性的集体控制。^[8]于是，学术自由被高举为一种理想，仿佛是在说：“你应当作为一个个体。你应当始终忠于自己认定的真理。”但任何围绕“为你构造的这个共同体究竟是谁”或“你的学科有怎样的历史”展开的社会学分析，却都被视为威胁个人科学自由的东西。我觉得，科赫兄弟兜售的正是这一套。

KA：与道德被抽空和去政治化这一点相关的是，权力的宣示并非无法审视，却经常被强行隐藏。部分原因在于，相互竞争的政治项目会居中塑造资助发生的方式。我曾在一篇题为〈编码印度教民族主义：Shalini Kantayya 与《编码偏见》〉的 *Medium* 博文中写过例子。^[9] 文章描述，《编码偏见》的制作者曾是美国印度教基金会的宣传代表人物，而该基金会与准军事组织 RSS 及 Narendra Modi 有联系。问题是，我不认为那部电影里90%的人都在喊“印度教民族主义万岁”；我觉得，他们真的毫不知情。他们被邀请出演一部笼统地反对大型科技企业的电影，就答应了。但一旦我们开始理解，美国80%的 H1B 签证流向印度，其中大多数又流向婆罗门，我们就能看到这种种姓主义如何在硅谷得到再生产。甚至参与解雇 Timnit 的一位主管也是婆罗门。而这又与一种反黑人性相联系。各种带有法西斯性质的政治项目，与这个资助领域彼此相连。但我们不可能全都知道一切，也不可能对一切都有足够的社会背景知识。比起巴西博索纳罗治下发生的事情，我对印度教民族主义、对这里的美国黑人政治，稍微多知道一点。而这些差异，也是遮蔽眼前权力动态的部分原因。许多决定是在紧闭的房间里作出的；要理解最后是谁进入了这些房间，本身就是一项漫长的历史研究。

AH：听着这些关于学术界去政治化的评论，我也想到 Khadijah 为 *Logic* 某一期写的前言。她提到，曼哈顿计划的科学家曾写信给杜鲁门，恳求不要投下原子弹。^[10] AI 科学家发现自己陷在一种颇为无措的政治空间里。我曾在谷歌内部听过 Geoff Hinton 的演讲，他是机器学习或 AI 的教父之一。AI，或者说深度学习，之所以重新兴起，是因为他们碰巧从20世纪80年代起就一直琢磨同一类技术，最后终于有了足够多的数据，让它真正奏效。但随后，这些研究科学家突然获得了极其巨大的权力，却不知道该拿它怎么办。Hinton 回忆，赢得 ImageNet 竞赛后，他们开始收到惊人的报价。他说，有一年他们在一家赌场参加 NeurIPS，别人为他与 Alex Krizhevsky、Ilya Sutskever 共同创办的公司开出数以百万、千万美元计的收购价；Sutskever 后来又共同创办了 OpenAI。

在那个时刻之前，这些人并没有形成政治分析。随后，他们开发的东西成了人人谈论的热门话题，数百万、甚至数十亿美元向他们砸来；这意味着，他们发展那种分析的动力反而更少了。

不过，那是计算机科学家的一种奇怪特性。我是社会学家，但即便社会学这个表面上与分析不平等有关的学科，其训练方式实际上也与改善不平等没多大关系。在社会学家每年的大型会议上，总会反复出现一个极其脱离现实的问题：你究竟应不应该参与某种政治。最糟的是，学科内有许多人根本就是彻头彻尾的右翼白人至上主义者。但与此同时，也有很多善良的自由派，只是非常高兴特朗普终于走了，觉得现在只要回到奥巴马执政时我们所做的事情就好。

所以，现在身处这个空间十分奇怪，因为这些活动完全无关痛痒。这些人原本只是好科学家，后来却被抛到权力位置上；这样的过程意味着，他们更有动力把这些事情在世界上可能产生的一切政治效果剥除干净。

KA：就实际情形而言，曼哈顿计划的科学家并不把自己看成政治行动者。我想，我在那篇 *Logic* 社论中想表达的一部分，就是那个小笑话：技术资本主义竟让没学过人文学科的人掌管人类。因此，去政治化就是把人文学科切除出去。同时，Frank Pasquale 在那本名义上谈机器人、其实并不只是谈机器人的书《机器人学新法则》中，也提出了这样一个论点：^[11]AI 正在取代人类的专业知识。危险在于，Zoom 会挤掉中间环节，我们将不再有大学，等等。这意味着，我们需要重建自己的机构、公民社会与国家；随着新的财富集中为这些新资本形式，它们已经被金融化、碎片化、拆散了。

但我之所以在那篇社论里特别呼唤黑人研究领域的人加入，并试图换一种方向，而不是继续点名批评那些道德空洞、让我发疯的人，是因为我知道，他们懂得比我多。黑人研究学者特别比我更了解的，是作为人意味着什么，以及启蒙思想有什么问题。因为当我想到道德上的明确界线，我就想到孟德斯鸠说“不要再施行酷刑”，以及“我们该怎样组织这些不同的政治共同体”。但与此同时，那种启蒙理性主义，那种“只做科学家”的观念，却以殖民主义和大规模剥夺为前提。它不只以 IBM 在大屠杀期间优化犹太人的运输为前提，也以更早发生的赫雷罗人和纳马人大屠杀为前提。

因此，对我而言，问题在于必须引入社会背景，必须引入人文学科。但与此同时，我们所有人都身处某种背景之中，我们思考所用的工具，已经曾把一些人排除在“人”这个范畴之外。连我们自己的想象力，也被它限制了。我并不觉得自己已经拥有理清这一切的全部工具，但我知道，它不可能只是个人的事情。我认为，如果我们只走个人路线，就会陷入麻烦。但我也不知道，集体行动究竟意味着什么。我相信我们需要这种劳工组织工作，需要有意识地行动；然而，我们仍在这些机构中运作。我不认为大学明天就会消失，也不认为谷歌会消失。那么，我们该怎样在这个结构层面行动？

AH：这与组织问题以及政治左翼的表现形式密切相连。Khadijah，我想起你大概几个月前在推特上发的一串帖文，其中提到，革命共产党是你最初的政治归属之一。我对此有些共鸣。我19岁时的第一个政治归属是“新闻与书信”（News and Letters），一个小型托洛茨基主义团体，在 Raya Dunayevskaya 与 C.L.R. James 分道扬镳之前由他们领导。他们有一份报纸，如今你还能看到戴着报童帽的年轻男子在大学校园里兜售它。但他们确实有一些工人思想者，认真地为这些刊物投稿；这些刊物也确实是相当有效的政治组织媒介。这属于一个更大的问题：政治组织工作究竟是什么样子，而现在这个空间又如此碎片化。关于左翼组织的状

况，已经写了无数文章。在这个大规模平台化的时代，要弄清劳动组织或政治组织该怎么做，并不容易。

Dean Spade 在 *Social Text* 的一篇文章中，^[12]区分了几种行动主义策略。首先是游行与公开行动，但此外还有各种照护架构，由此引出对互助重要性的讨论。他那本更完整讨论互助的书，^[13]很大篇幅都在讲“怎样开会”，怎样让人承诺做事，比如怎样使用报名表。我觉得这太棒了，因为它很好地促使我们回过头来思考：如果从最基础的入门工作做起，这些网络会是什么样子。

JG: Alex，我觉得你指出了一种真实的张力：一边是一些极其个人化的要求，它们构成了你此前所称“声望经济”的重要部分——成为那个因为揭露坏企业而获得赞誉、得到演讲邀请和学术荣誉的人，等等；另一边则是政治组织工作的集体要求，而后者完全得不到同样的奖励。这必然与我们的机构生活如何组织有关。

LI: 我不知道，你说“机构”时，指的是大学的设置方式吗？你是在想这个吗？我对大学怀有极其矛盾的感情。一方面，这是从阅读 Moten 与 Harney 关于底层公地和逃逸规划的论述中学到的；^[14]他们认为，大学在某种意义上是制造治理知识的场所。另一方面，这也来自我教过的那些没有合法移民身份的本科生：他们来到这里，有了一个远离成长环境的空间，可以遇见新的人、新的想法，成为自己想成为的人。

所以，我想，我把大学看作雇主：我要试图改革我的雇主吗？还是要通过组织工作，争取在长远上对雇主拥有某种实质性的民主所有权，而我愿意为之奋斗的改革，应当是那些帮助我们想清楚这种所有权到底可能是什么，并建立实现它所需技能与关系的改革？

还有，刚才 Khadijah 谈“什么是人”的问题时，我也在想：我完全理解，你说黑人研究能够回应这个问题，指的是什么；但归根结底，我也不会把这个问题的解释权让给任何学者。我会从黑人研究及其他人所能教给我们的东西中获得启发，但我们也必须通过自己的组织工作，找到答案。也许我对组织工作有些理想化，但我觉得，它可以以这样一种方式开展：当人们受到共同生活这一难题的挑战、不得不设法弄清怎样彼此相处时，它能激发出他们所能形成的承诺中最好的部分。

KA: 我听到了你对学术界的矛盾感情。我想，在那篇 *Logic* 社论里，我部分想说的是：Audre Lorde 的那句话——你不能用主人的工具拆毁主人的房子——有非常多层含义。学术界并不拥有知识生产。因此，当我想到黑人研究时，我甚至不一定指那种特定的文本形式，也就是我们通过学术讲座、以一种近乎礼拜仪式的方式传播的东西。它可以以如此多的方式存在。

另一个张力，围绕着通过大学提供资助。作为一个始终身处边缘的人，我与大学的关系，也受到自己对他人责任的塑造。因为对我而言，某种集中化是必要的。如

今，去中心化很流行，但问题在于：我要怎样把资源交到人们手中？获得资助的人可以成为传递资金的渠道，并希望建立一种自己曾经没有得到的东西。我写完〈AI 伦理的道德崩塌〉时，^[15]Safiya Noble 开玩笑问：“你有医疗保险吗？”我说：“没有啊，老兄，我穷，是真的穷。”然后，他们帮了我。我从他们那里得到资助，而且你知道吗，他们就让我自己做事。给我资助的人——哥伦比亚大学、纽约大学、UCLA——从来没有谁问过我到底在干什么，从来没有。没人来检查我，没有截止日期，没有必须参加的会议，我从来不上 Zoom，从来不上。没人管我在做什么。当然，另一方面，我最终还得去找2023年的新资助。但我想说的是，人们可以通过大学，为其他人创造这样的条件。我觉得，Meredith 和 Safiya 展现了真正的照护责任，而且不只是对我。

在我愿意使用这个词的意义上，我也觉得自己拥有学术自由。我觉得自己可以提出一个项目，再向别人推介。例如，我可以批评 Ben [Tarnoff] 和 Moira [Weigel]，但他们让我在 *Logic* 按自己的想法做事，确实没有干预。这些机会是存在的。我想，Minderoo 资助了 UCLA 的孵化器，而我获得了一笔社会影响力资助，总共大概50,000美元。我得以资助全球六个项目、克什米尔的六个人，以及奥罗米亚的一个人；我也不向他们提出任何要求。我说：“这笔钱，你们想怎么用就怎么用。”“如果我能帮忙，我会帮，但你们不用汇报任何东西。”通常，*Logic* 每篇文章付给作者250至500美元，但我们给了2,000至4,000美元。我的想法，是把现金注入所有那些做着很棒的工作、却总是身无分文的人手里。4,000美元不会改变一生，但能给人一些帮助。我只是想给你们一些具体例子，说明资助可以是什么样子。当然，这里面同时也牵涉邪恶的事情，但我不知道该怎么处理。

AH：有几件事，我想继续谈。当我说到“声望经济”，说到某些人是演讲活动必须“招募”的对象时，并不是说，他们接受演讲邀请就是错的。我自己也接受过很多邀请。更准确地说，学术界的激励机制高度围绕演讲、发表等事情进行优化。最糟的情况是，有些人在公众面前十分活跃，却在私下传递警示的信息网络里被公认为施虐者；他们从年轻人那里抽取各种劳动，同时使用解放、AI 伦理或其他什么语言——词汇随你挑。这并不是在服务于 Khadijah 所说的对人的责任。究竟什么叫对 X 的责任、对人的责任，是我不断返回的一个分析问题。我说“对 X 的责任”，是因为 X 会随着你所处位置而变化。例如，在谷歌，它是对用户的责任——这就是其表述方式，人变成了用户。在非营利组织中，它是对资助者或其他人的责任。说了这么多，我想表达的是，一个人的主体位置非常重要；而要抵抗这些激励机制，需要付出极大的努力，因为你所在机构的所有其他人，都在紧紧盯着你、给你施压。

我开始读 Sara Ahmed 的《投诉！》。^[16]她讨论的一点是，在机构内部提出投诉，会非常清楚地揭示那些激励机制是什么，以及机构内部如何运作。目前，关于人们如何被卷入声望与发表的浪潮，还没有真正的学术分析。

LI：我进入学术界，是因为觉得科技行业里没有真正的空间，让我从事与自身价值观一致的工作；我当时不知道，那其实只是资本主义异化的一个特征。因此，在学术界，你还会被另一种希望裹挟：“哦，我想做批判性工作，想说那些在公司里不能说的话。”但目的是什么呢？那种“说”也未必带来任何结果。我想说的是，有些人在追逐声望，有些人在追逐理想主义的希望，但无论哪种情况，赢的都是资本。

JG：说到为资本服务的理想主义，你们刚才提到了给 FAccT 组织委员会写信。FAccT 在 AI 伦理/科技政策领域似乎占据着重要的机构位置，因为它有能力决定什么被纳入“公平”领域的议程，并在一定程度上决定更广义的 AI 伦理议程；它还能够生成关系网络，促进产业与学术界互动。在我看来，要求 FAccT 表明立场，可能是一项非常有意义的行动。推动这项行动的动力从何而来？

LI：接下来这些话，我可能全都会删掉。

给 FAccT 写信的起因，是一位与会者在公开讨论会上，对 Christo Wilson 的论文非常愤怒。那篇论文基本上让一家名为“pymetrics”的公司，自行界定对其招聘算法进行公平性审计的条件。^[17]后来，这家公司到处拿那篇论文作宣传，作为其接受独立审计的例子，好让地方市议会和监管者认为，这套算法已经可以投入使用。撰写那封信的过程，随后又扩展为对其他问题的讨论。Khadijah 从一开始就参与了，Alex 也参加了那些讨论。

要求 FAccT 表明立场，或者看清它不会公开表明立场，都是有意义的行动。我很好奇其他人怎么看。像很多组织工作一样，这是在做中学。因此，我和 Alex、我和 Khadijah，以及签署那封信的人组成的网络之间，进行了很多谈话；一些没有签名、但很高兴这件事正在发生的人，也参与了。我们会问：“对，提出要求有意义吗？针对这件具体事情提出要求，有意义吗？”这迫使我们在一定程度上澄清对这个机构究竟是什么的分析，并使其或许能够公开表达。

但坦白说，我也只是希望，我们能开一次集体会议，把这些事谈透，因为整个过程感觉耗时太长，需要太多一对一、幕后渠道里的谈话。信任和时间的问题都太紧迫了。

KA：FAccT 有意思的一点是，我从来没去过 FAccT。我在推特上带着厌恶旁观它，已经很久了。

对我而言，那封信很完整地体现了“谁是我们的‘共同体’”这个问题。参与者之间的立场鸿沟——“去他妈的警察，警察全是混蛋（ACAB）”这一方，与“我就是为他们工作的”那一方之间的鸿沟——在我看来无法弥合。但让我特别注意的，是 FAccT 的组织能力。从机构资源和线下活动来说，我们任何由自己组成的网络，都不可能复制那种能力。而在数字空间里，它们维持的霸权，对我而言，也体

现了这种矛盾感的程度，以及为什么组织学者如此困难。好吧，大家都疲惫不堪；在行政操作上，我也承认办一场会议多少有些困难。但我真的很挣扎。我觉得，到某个程度，你根本不可能再推动那种人向前，他们就是会拒绝。特别是，我厌恶 FAccT 的许多原因之一，是联系他们后发现，执行委员会的一位成员竟然是阿勒格尼县家庭筛查工具的开发之一；对我来说，这实在太过分了。

所以我问：“为什么我们不办自己的会议？”整件事本来就建立在无偿劳动之上，而开会，尤其给人一种我们真能做成的感觉。哪怕只办三场工作坊，叫它——我不知道——另类 FAccT 会议，或者前 FAccT 退出者会议。但去做这件事的动力在哪里？我觉得，如果我们规定大家必须做，他们就会做。但要自发掀起一股浪潮，这种想法似乎完全处在集体想象之外。

AH：我会问的一个问题是：为什么还要围绕 FAccT 开展组织工作？Lilly 和我为此谈过很多次。这些人只会固守阵地。但也许，迫使他们亮明底牌，仍然值得？

最后，Lilly 很好地说服了我。但我也觉得，围绕 FAccT，这正在变成一场没完没了的对话。2018年，我参加了纽约的第一届线下 FAccT。记得当时有人站起来，直接说：“这他妈的是什么？我们为什么要在乎？”他们说，这些讨论基本上都无关实质：“你们只是在想怎样以公平的方式监视人。”

LI：FAccT 那封信有一点很重要：通过界定 FAccT 到底有什么问题，并让人们签名公开说这件事是错的，围绕它展开的讨论帮助我们澄清了这一点——如果要像 Khadijah 所说的那样，另办一场会议，我们就必须明确，它应当承担哪些承诺。

另一个在我脑子里不断盘旋的想法是：已经非常清楚，FAccT 的承诺是维持一个大帐篷，或者用他们反复说的话，叫“维持一张讨论桌”。谈到 PLSC 时，他们也用这个词。但这反映了一种幻想：只要让思想自由交换，正确的事情就会发生。

JG：感谢大家为这场对话带来的细致区分。退出是个艰难的选择，但你们确实展现了它与学术生活复杂性之间的关系。研究者和学者始终在协商自己与联盟和共同体的关系、金钱和声望的政治、机构要求与学术自由，以及我们最终究竟对谁负有责任。你们真正提醒了我们，眼前究竟有什么利害攸关；对此，我感激不尽。

参考文献

Abdurahman, J. Khadijah. ‘On the Moral Collapse of AI Ethics’ [〈AI伦理的道德崩塌〉]，Medium blog, 8 December 2020, <https://upfromthecracks.medium.com/on-the-moral-collapse-of-ai-ethics-791cbc7df872>.

——. ‘Encoding Hindutva: Shalini Kantayya and Coded Bias’ [〈编码印度教民族主义：Shalini Kantayya 与《编码偏见》〉]，Medium blog, 2 June 2021, <https://upfromthecracks.medium.com/encoding-hindutva-shalini-kantayya-and-coded-bias-721fe04f225f>.

——. ‘Holding to Account: Safiya Umoja Noble and Meredith Whittaker on Duties of Care and Resistance to Big Tech’ [〈追究责任：Safiya Umoja Noble与Meredith Whittaker谈照护责任和抵抗大型科技企业〉]，*Logic: Beacons* 15 (2022), <https://logicmag.io/beacons/holding-to-account-safiya-umoja-noble-and-meredith-whittaker/>.

——. ‘A Body of Work That Cannot be Ignored’ [〈不可忽视的一系列工作〉]，*Logic: Beacons* 15 (2022), <https://logicmag.io/beacons/a-body-of-work-that-cannot-be-ignored/>.

Ahmed, Sara. *Complaint!* [《投诉!》]，Durham: Duke University Press, 2021.

Dickenson, Tim. ‘Inside the Koch Brothers Toxic Empire’ [〈科赫兄弟有毒帝国内幕〉]，*Rolling Stone*, 24 September 2014, <https://www.rollingstone.com/politics/politics-news/inside-the-koch-brothers-toxic-empire-164403/>.

Harney, Stefano and Moten, Fred. *The Undercommons: Fugitive Planning & Black Study* [《底层公地：逃逸规划与黑人研究》]，London: Minor Compositions, 2013.

Meyerhff, Eli, Johnson, Elizabeth, and Braun, Bruce. ‘Time and the university’ [〈时间与大学〉]，*ACME: An International E-journal for Critical Geographies* 10(3) (2011): 483–507. [译注：作者姓氏Meyerhff沿用原书拼写。]

Pasquale, Frank. *New Laws of Robotics: Defending Human Expertise in the Age of AI* [《机器人学新法则：在AI时代捍卫人类专业知识》]，Cambridge: Harvard University Press, 2020. [译注：出版地Cambridge为原书拼写。]

Phan, Thao, Goldenfein, Jake Mann, Monique, and Kuch, Declan. ‘Economies of Virtue: The Circulation of “Ethics” in Big Tech’ [〈美德经济：“伦理”在大型科技企业中的流通〉]，*Science as Culture* 31 (2022): 121–35. [译注：原书作者著录在Jake与Mann之间缺少分隔标点，此处保留。]

Spade, Dean. ‘Solidarity Not Charity: Mutual Aid for Mobilization and Survival’ [〈团结而非慈善：为动员与生存而互助〉]，*Social Text* 142(38) (2020): 131–51.

——. *Mutual Aid: Building Solidarity During This Crisis (and the Next)* [《互助：在这场危机及下一场危机中建立团结》]，London: Verso, 2020.

Wilson, Christo, Ghosh, Avjit, Jiang, Shan, Mislove, Alan, Baker, Lewis, Szary, Janelle, Trindel, Kelly, and Polli, Frida. ‘Building and Auditing Fair Algorithms: A Case Study in Candidate Screening’ [〈构建与审计公平算法：候选人筛选案例研究〉]，*Proceedings of the FAccT Conference*, March 1–10, 2021, Virtual Event, Canada, 1–10 March. 2021. 会议信息：《FAccT会议论文集》，加拿大线上活动，2021年3月1—10日。

原注

[1] Alex Hanna 此后已离开谷歌，现任分布式人工智能研究所（Distributed AI Research Institute, DAIR）研究主任。[译注：“现任”指原书编定时。] 返回

[2] 参见本书《“公开的秘密”》。返回

[3] Timnit Gebru 是一位知名计算机科学家，曾任谷歌研究部门伦理AI团队共同负责人。2020年12月，谷歌宣布她已辞职；Gebru 否认这一说法，称自己是在围绕一篇学术论文发生争议

后被解雇的。在那篇论文中，她与其他同事对大型语言模型提出了伦理风险方面的质疑。
返回

- [4] 关于科赫兄弟与政治洗白的更多讨论，参见 Tim Dickenson, ‘Inside the Koch Brothers Toxic Empire’ [〈科赫兄弟有毒帝国内幕〉]，*Rolling Stone*, 24 September 2014, <https://www.rollingstone.com/politics/politics-news/inside-the-koch-brothers-toxic-empire-164403/>. 返回
- [5] ‘Holding to Account: Safiya Umoja Noble and Meredith Whittaker on Duties of Care and Resistance to Big Tech’ [〈追究责任：Safiya Umoja Noble与Meredith Whittaker谈照护责任和抵抗大型科技企业〉]，*Logic: Beacons* 15 (2022), <https://logicmag.io/beacons/holding-to-account-safiya-umoja-noble-and-meredith-whittaker/>. 返回
- [6] Thao Phan et al., ‘Economies of Virtue: The Circulation of “Ethics” in Big Tech’ [〈美德经济：“伦理”在大型科技企业中的流通〉]，*Science as Culture* 31 (2022): 121. 返回
- [7] FAccT 是美国计算机协会“公平、问责与透明”网络（Fairness Accountability and Transparency）的缩写，该网络每年举办名为 ACM FAccT 的会议。返回
- [8] 此处指《1940年学术自由与终身教职原则声明》（1940 Statement of Principles on Academic Freedom and Tenure），<https://www.aaup.org/report/1940-statement-principles-academic-freedom-and-tenure>; Eli Meyerhff et al., ‘Time and the University’ [〈时间与大学〉]，*ACME: An International E-journal for Critical Geographies* 10(3) (2011): 483。[译注：Meyerhff沿用原书拼写。] 返回
- [9] J. Khadijah Abdurahman, ‘Encoding Hindutva: Shalini Kantayya and Coded Bias’ [〈编码印度教民族主义：Shalini Kantayya与《编码偏见》〉]，*Medium blog*, 2 June 2021, <https://upfromthecracks.medium.com/encoding-hindutva-shalini-kantayya-and-coded-bias-721fe04f225f>. 返回
- [10] J. Khadijah Abdurahman, ‘A Body of Work That Cannot be Ignored’ [〈不可忽视的一系列工作〉]，*Logic: Beacons* 15 (2022) <https://logicmag.io/beacons/a-body-of-work-that-cannot-be-ignored/>. 返回
- [11] Frank Pasquale, *New Laws of Robotics: Defending Human Expertise in the Age of AI* [《机器人学新法则：在AI时代捍卫人类专业知识》]，Cambridge: Harvard University Press, 2020. 返回
- [12] Dean Spade, ‘Solidarity Not Charity: Mutual Aid for Mobilization and Survival’ [〈团结而非慈善：为动员与生存而互助〉]，*Social Text* 142(38), (2020): 131. 返回
- [13] Dean Spade, *Mutual Aid: Building Solidarity During This Crisis (and the Next)* [《互助：在这场危机及下一场危机中建立团结》]，London: Verso, 2020. 返回
- [14] Stefano Harney and Fred Moten, *The Undercommons: Fugitive Planning & Black Study* [《底层公地：逃逸规划与黑人研究》]，London: Minor Compositions, 2013. 返回
- [15] J. Khadijah Abdurahman, ‘On the Moral Collapse of AI Ethics’ [〈AI伦理的道德崩塌〉]，*Medium blog*, 8 December 2020, <https://upfromthecracks.medium.com/on-the-moral-collapse-of-ai-ethics-791cbc7df872>. 返回

-
- [16] Sara Ahmed, *Complaint!* [《投诉!》], Durham: Duke University Press, 2021. [返回](#)
- [17] Christo Wilson et al., ‘Building and Auditing Fair Algorithms: A Case Study in Candidate Screening’ [《构建与审计公平算法：候选人筛选案例研究》], Proceedings of the FAccT Conference, Virtual Event, Canada, 1–10 March 2021. 会议信息：《FAccT会议论文集》，加拿大线上活动，2021年3月1—10日。 [返回](#)

致谢

许多出色的人帮助我们让这本文集得以成形。感谢才华横溢的作者们，不仅贡献了时间，也慷慨投入了实验性的“集体评审”过程。感谢Geert Lovink对本项目的耐心与支持，以及他对于颠覆学术实践中束缚性规范的出版方式所抱持的信念与远见。感谢*Science as Culture*（《科学作为文化》）的编辑和审稿人，尤其是Kean Birch与Kelly Bronson；他们编辑了“大型科技企业论坛”（Forum on Big Tech），我们最初关于“美德经济”的文章便发表于该论坛。感谢Elena Gomez迅速而敏锐的编辑工作。感谢参加2021年7月首届“美德经济”工作坊的报告者：Angela Daly、Sarah Pink、Brett Neilson、Mark Andrejevic、Julia Powles、Sy Taffel、Winifred Poster、Tsvetelina Hristova、Michael Richardson、Tanja Dreher和Ed Santow。感谢Emma Kowal给予支持与指导，帮助我们申请澳大利亚社会科学院工作坊计划资助。最后，感谢在我们探索这一棘手主题时，慷慨地与我们交谈并提供更多帮助的众多交流者，尤其是Seth Lazar、Kate Devitt、Christine Parker和Julia Powles。

撰稿人简介

J. Khadijah Abdurahman（代词：they/them/任何代词）是一位废除主义者，研究重点是纽约市儿童福利系统中的预测性风险建模，以及非洲之角的技术。他们即将担任 *Logic(s)* 杂志主编，是UCLA C2I2研究员，也是哥伦比亚大学INCITE中心公共利益技术项目We Be Imagining的创办人。Khadijah与华盛顿大学人本设计与工程系（HCDE）Sucheta Ghoshal主持的Inquilab项目共同创办了Otherwise School：“反法西斯的工具与技术”（Tools and Techniques of Counter-Fascism）。

Laura Bedford是迪肯大学犯罪学讲师，致力于推进绿色犯罪学、环境犯罪、抵抗和行动主义相关知识的发展。她尤其关注将占据霸权地位的犯罪学理论与刑事司法实践，移用到英语西方世界之外时所产生的困难，并对这一过程提出问题。

Corinne Cath是一位研究互联网基础设施政治的文化人类学家。他们于2021年在牛津大学完成博士学位，现为剑桥大学艺术、社会科学与人文研究中心（CRASSH）下属Minderoo技术与民主中心的联属研究员。

Angela Daly是一位研究（数字）技术的法律社会学学者，现任苏格兰邓迪大学法律与技术教授，由Leverhulme法医学研究中心与法学院联合聘任。她担任苏格兰政府“释放数据价值”独立专家组主席，并是苏格兰政府警务新兴技术独立咨询小组成员，负责法律与伦理工作领域。她还是意大利米兰大学信息社会法律中心研究员，在那里开展人脸识别监管研究。

Jake Goldenfein是一位法律与技术学者，重点研究平台监管、监控与自动化决策治理。他是墨尔本大学墨尔本法学院高级讲师，也是澳大利亚研究理事会自动化决策与社会卓越研究中心的关联研究员。

Tsvetelina Hristova是西悉尼大学文化与社会研究所博士后研究员。她的工作位于媒体研究、人类学和科学技术研究（STS）的交叉领域，所发表的研究聚焦组织与基础设施的社会物质实践的政治，以及这些实践与劳动管理、移民和政治经济的关系。

Alex Hanna是分布式人工智能研究所（Distributed AI Research Institute, DAIR）研究主任。她受过社会学训练，工作聚焦新计算技术所使用的数据，以及这些数据如何加剧种族、性别与阶级不平等。

Lilly Irani是加州大学圣迭戈分校传播与科学研究副教授，著有 *Chasing Innovation: Making Entrepreneurial Citizens in Modern India*（《追逐创新：现代印度创业型公民的塑造》，Princeton University Press, 2019），以及与Jesse

Marx合著的*Redacted*（《删改》，Taller California, 2021）。她还与科技劳动者联盟，以及圣迭戈透明且负责任地使用监控技术联盟（Transparent and Responsible Use of Surveillance Technology, TRUST）一同开展组织工作，并担任圣迭戈出租车劳动者联合会理事。

Os Keyes是华盛顿大学博士候选人，也是首批Ada Lovelace研究员之一。他们研究AI中的性别与残障问题，以及跨性别医学与科学的历史。

Declan Kuch是西悉尼大学文化与社会研究所可持续性与伦理领域的校长研究员。他的研究聚焦民主与新兴科学技术之间的互动。

Monique Mann是迪肯大学犯罪学高级讲师，也是该校Alfred Deakin公民身份与全球化研究所成员。Monique研究新技术的社会正义意涵，以改善对这些技术的监管。

Liam Magee是西悉尼大学文化与社会研究所副教授兼社会合作主任。Liam的主要研究兴趣，是将社会研究方法和信息技术应用于城市发展与可持续性领域。

Rodrigo Ochigame是荷兰莱顿大学文化人类学与发展社会学研究所助理教授。

Sarah Pink是莫纳什大学教授兼新兴技术研究实验室主任。Sarah是未来与设计人类学家，也是电影制作人；近期作品包括著作*Emerging Technologies / Life at the Edge of the Future*（《新兴技术／未来边缘的生活》，2022）与纪录片*Digital Energy Futures*（《数字能源未来》，2022）。

Thao Phan是澳大利亚研究理事会自动化决策与社会卓越研究中心、莫纳什大学新兴技术研究实验室的研究员。她是一位女性主义科学技术研究（STS）学者，研究算法文化中性别化与种族化的面向。

Michael Richardson是新南威尔士大学（UNSW）媒体学副教授，并共同主持该校“媒体未来中心”。他的研究探讨新兴技术如何塑造文化与权力，又如何受二者塑造。他的下一本著作*Nonhuman Witnessing: War, Data, and Ecology After the End of the World*（《非人见证：世界终结之后的战争、数据与生态》），即将由Duke University Press出版。

Jathan Sadowski是莫纳什大学新兴技术研究实验室高级研究员，正在以澳大利亚研究理事会DECRA研究员项目资助，研究保险技术的政治经济。他著有*Too Smart: How Digital Capitalism is Extracting Data, Controlling Our Lives, and Taking Over the World*（《过度智能：数字资本主义如何提取数据、控制我们的生活并接管世界》，The MIT Press）。

Sy Taffel是奥特亚罗瓦—新西兰梅西大学媒体研究高级讲师、政治生态研究中心共同主任，著有*Digital Media Ecologies*（《数字媒体生态》，Bloomsbury，2019）。他的研究聚焦政治生态学与数字技术的交叉领域。

Meredith Whittaker是Signal总裁，也是AI Now研究所的共同创办人兼首席顾问，曾任美国联邦贸易委员会AI高级顾问。

理论按需 第46号

美德经济：人工智能领域中“伦理”的流通

编者：Thao Phan、Jake Goldenfein、Declan Kuch、Monique Mann

AI伦理从未远离它试图批判的产业。它最初旨在把公平、问责、透明等价值带给大型科技企业及其产品，但科技企业的公关举措与AI伦理资助之间，界线从未清晰。在实践中，AI伦理如今成了收编批评者、促成监管俘获的手段。企业借它制造正当性，进一步积累价值。结果是，“伦理”已成为一种高价值的工业商品，而AI伦理就是铸造这种商品的工厂。

本选集集体回应了伦理被物化为商品形式的过程。它探讨产业参与“伦理AI”研究，如何创造出一种新的“美德经济”：一个庞大的行动者网络，分布于产业、公民社会和大学，把伦理作为服务与产品加以生产和流通。撰稿人既提供对这一经济的批判视角，也分享亲身经历。议题广泛，包括在产业资助空间中工作的矛盾与个人困境；国防、人脸识别、标准制定等领域的AI伦理案例；对漂绿、制造信任等手法的批判评估；以及公开发声、组织反对和退出等直接行动的风险与实际操作。这些文章共同说出了困扰AI伦理的收编、俘获和共谋难题，并使界定这一领域的网络与流通关系显现出来。

作者：Corinne Cath、Os Keyes；Sarah Pink；Rodrigo Ochigame；Sy Taffel、Laura Bedford、Monique Mann；Angela Daly；Tsvetelina Hristova、Liam Magee；Michael Richardson；Jake Goldenfein、Lilly Irani、J. Khadijah Abdurahman、Alex Hanna；Jathan Sadowski、Thao Phan、Meredith Whittaker。

按需印刷。ISBN：9789492302960

网络文化研究所