arXiv:2602.01824v1 [cs.HC] 2 Feb 2026

风险、数据与对齐：让肯尼亚的信用评分发挥作用

Daniel Mwesigwa
dm663@cornell.edu
康奈尔大学 伊萨卡，纽约州，美国

Steven J. Jackson
sjj54@cornell.edu 康奈尔大学 伊萨卡，纽约州，美国

Christopher Csikszentmihalyi
cpc83@cornell.edu 康奈尔大学，美国纽约州伊萨卡

摘要

信用评分日益成为数据与人工智能治理中至关重要且备受争议的领域，常被描述为一种中立、客观的方法，用于在多元的经济与政治情境中评估风险。我们基于对肯尼亚内罗毕信用评分实践开展的为期九个月的民族志研究，考察了数字借贷中数据科学的社会技术工作与制度性工作。区域内成熟的电信公司和银行正利用专有数据开发数字贷款产品，与此同时，算法信用评分也在新的行动者、技术和不断变化的法规推动下发生转变。研究发现，从业者如何通过技术和法律变通方法构建替代数据，基于多种解读来界定风险，并借助技术和政治手段协商模型性能。我们认为，算法信用评分是通过持续不断的对齐工作实现的；这种工作在持续的不确定性条件下稳定风险，并呈现出认识论、建模和情境等形式。通过拓展人机交互领域关于对齐的研究，我们展示了对齐如何作为一种双向转译发挥作用：模型被塑造得“适用于世界”，而那些世界也被重塑得“适用于模型”。

CCS概念

•**以人为中心的计算** → **人机交互实证研究**；理论、模型与概念。

关键词

算法信用评分、数字借贷、替代数据、民族志、数据科学、机器学习实践、肯尼亚

ACM参考文献格式：

Daniel Mwesigwa、Steven J. Jackson与Christopher Csikszentmihalyi。2026。风险、数据与对齐：让肯尼亚的信用评分发挥作用。载于*2026*年*CHI*计算机系统中的人为因素会议论文集（*CHI '26*），*2026*年*4*月*13*日至*17*日，巴塞罗那，西班牙。ACM，纽约，NY，美国，19 页。https://doi.org/10.1145/3772318.3790924

1 引言

在肯尼亚内罗毕，随着消费者对信贷的需求不断增长，金融科技公司也努力扩大市场覆盖范围，新的算法信用评分模型应运而生。这些模型通常采用人工智能（AI）、机器学习（ML）和数据科学技术开发 [81, 88, 90, 97, 110, 116]。肯尼亚在非洲数字借贷领域处于领先地位：世界银行《2025年全球普惠金融数据库报告》 [58] 显示，32%的肯尼亚成年人会借款其中，25%的人只通过移动货币服务提供商借款；86%的借款人使用移动货币支付学费和日常消费等开支。尽管数字金融服务的普及程度显著提高，肯尼亚的金融科技公司仍被指控收取高利率（年化利率超过100%）、采取公开羞辱等激进的债务催收手段，以及侵犯消费者隐私 [56, 81, 89]。借款人则采取了各种做法来应对信贷义务，包括更换 SIM（身份）卡，或从其他渠道借款以偿还数字贷款 [90, 117, 120]。由此，数字贷款违约率不断上升，最高曾达到40% [77]。数字信贷提供商，包括电信公司、银行和初创企业，如今正通过加倍投入算法信用评分等技术，并纳入有关当地信贷文化和消费者行为模式的新数据来源，以应对这些违约和“风险前沿”。这些以技术为中心的应对措施看似新颖，但二十多年来一直被视为金融普惠工具包的一部分。例如，人们早已探索移动货币和通话时长交易日志等数据来源，评估其提升算法信用评分的潜力，尤其是在传统信用征信机构的覆盖范围不完整或有限的情况下 [122]。发生变化的是新型机器学习技术和数据集的使用规模，以及受其影响的消费者数量。

信用评分是算法风险管理的典型例子，凸显了算法如何在高风险决策领域中开发并嵌入其中 [57, 63]。事实上，算法技术和机器学习技术用于预测的场景并不局限于信用风险评估，还包括刑事司法 [7], 、社会福利 [105], 和医疗保健 [23]。人机交互（HCI）及相关领域中越来越多的研究，为自动化决策系统及其社会影响提供了宝贵见解。例如，Barocas 和 Selbst [12] 指出，大数据的广泛能力如何利用性别、种族和宗教等受保护属性的代理特征，通过自动化决策系统再现历史偏见与社会不平等。由于公司严密保护其算法和工具，将其视为专有资产，因此难以调查这些技术背后的全部假设、设计选择和价值观 [48]；不过，人工智能审计和以社区为中心的研究方法已揭示真实用户的关切，并为问责开辟了切实可行的路径 [29, 60, 64, 114]。其他研究则立足于实验室环境中的工作场所民族志传统，探讨塑造信用评分的核心组织实践与协作实践，以及这些实践如何与不同的社会世界相遇，并由其塑造。这些研究基于实地观察，记录了信任 [92] 和透明度 [38] 等价值观如何在数据科学的日常工作中得到界定和落实，

本作品采用知识共享署名 4.0 国际许可协议。*CHI '26*，西班牙巴塞罗那
©2026 版权所有归所有者/作者所有。ACM ISBN
979-8-4007-2278-3/2026/04https://doi.org/10.1145/3772318.3790924

并如何在组织情境、不断变化的政策要求以及更广泛的社会利益与价值观之间实现对齐。

除了这些通常以西方为中心的地点和背景外，对全球多数世界许多地区的数据科学实践开展的组织研究，范围一直有限，并由一种以扩大数字信贷（及其他金融服务）覆盖面为导向的“包容主义”范式主导；这种范式立足于一种认知，即这些地区数据（和信贷）匮乏 [6, 10, 85]。例如，Shema [110]对卢旺达一家领先电信公司的隐私保护型信用评分效能进行的研究报告称，与使用大量侵犯隐私的消费者数据相比，使用有限数据（即话费充值）也能达到相当的性能指标。在另一种背景下，Speakman 等人 [116] 在一个缺乏相关训练数据的东非国家推出了一款基于手机的信贷产品，并借助该地区另一个数字借贷生态系统已充分发展的国家的数据集。这些干预措施为技术上哪些做法有效提供了有益见解，却没有说明这类系统在实践中如何构建，尤其没有揭示肯尼亚信用评分更广泛且独特的动态：技术竞争的不均衡，包括 Safaricom 对电信和金融科技数据近乎垄断（95+%） [74]；用户体验中体现的独特信贷文化，反映出根深蒂固的信任与风险共担实践 [8, 60]；以及不断变化的政策和监管环境，其围绕消费者数据保护和一类新兴的“公私”市场干预措施展开 [119]。

基于2025年二月至十一月开展的大规模民族志田野调查，我们追踪了数据科学和产品团队为肯尼亚市场设计和构建新型信用评分系统时所采用的技术、组织与文化实践，也考察了借款人、受益者和社区为使用这些系统并管理自身与其关系而采取的实践。本文主要报告初创企业、电信公司和银行的实践，探究它们如何将抽象的信用风险概念转化为信用评分工具与系统，并使其产生重大的市场影响和现实影响。

本文认为，在肯尼亚，算法信用评分是通过持续不断的对齐行动实现的，这些行动在长期不确定的条件下稳定了风险。我们指出，风险与不确定性并非彼此独立的现象，而是深度交织：信用评分模型声称要衡量的风险，始终受到三种不确定性的干扰，分别关乎哪些信息能够被知晓、模型是否如其所宣称的那样有效，以及人们对模型预测行使的自由裁量权。本文借鉴Fujimura[43] 和Dourish [37] 关于对齐的理论，展示技术问题如何在组织与社会世界中变得“可处理”，并说明尽管风险与不确定性交织在一起，信用评分仍是如何得以运作的。不同于人工智能安全论述中呈现的对齐观 [2, 51]，我们认为，对齐是一种双向转译：既让模型适应各种世界，也让各种世界适应模型。这种双重效应是信用评分发挥效用的核心，同时也深刻影响着其在隐私、监控和金融普惠这些关键且充满争议的领域中所带来的社会影响，无论好坏。

我们首先回顾人机交互（HCI）及相关领域关于风险、不确定性和对齐的研究，以及这些概念如何在算法信用评分等高风险决策领域中相互作用。随后，我们将介绍我们的研究场地和方法。研究结果记录了从业者如何利用非传统数据来源构建替代数据，展示了风险如何以多种方式被建构，并强调了协商模型表现的过程。我们的讨论认为，替代数据的形成具有施为性，会再生产出多种不同的风险影响；同时考察风险与不确定性的纠缠如何暴露预测的局限性；并讨论对齐的重要作用、其运作方式以及它服务于谁。最后，我们提出一些关于政策和社会技术研究的问题。

2 风险、不确定性与对齐

自动化决策系统在高风险领域的广泛应用，引发了人机交互和以人为中心的机器学习领域的广泛关注与担忧 [16, 68, 73, 75, 98, 108, 121]。人机交互及相关领域研究了预测系统中“风险”被界定和付诸实践的多种情境，包括信用评分 [48, 92], 、社会工作与福利 [27, 54, 105], 、刑事司法 [7, 53, 76], 以及医疗保健 [23, 49, 102]。人机交互领域近期的研究质疑了风险的认识论基础，探讨风险如何在预测模型中被界定和付诸实践，以及支撑风险概念的种种要求、组织实践和社会世界。例如，Saxena 等人 [106] 对美国儿童福利个案记录进行的计算叙事分析发现，福利个案中的问题源于个案工作者未能指出行政数据中的错误，以及个案工作者与父母之间不对称的权力关系；这掩盖了超出获准使用的“静态构念”范围的系统性和时间性风险因素。Wang 等人 [121] 指出，预测优化系统会将复杂的社会和组织问题简化为技术优化任务，往往反而加剧其试图缓解的风险。这一日益发展的研究方向与科学史和科学社会学的奠基性研究相呼应；后者表明，将风险认定为可计算并非中立之举，而是历经漫长而持久的分类与量化过程才得以实现 [18, 31, 94]。经济学家 Frank H. Knight [59, 在其对风险与不确定性的著名阐释中，将风险界定为可衡量的事物，“既可以通过先验计算，也可以根据过去经验的统计数据来衡量”；而不确定性则无法衡量，因为它受到无法预见的（未来）因素影响。因此，商业及相关领域的决策并非只是管理可量化的风险，而是在不可知的未来中采取行动 [59], ，这使得不确定性管理（或“驯服偶然性” [46]）成为成败攸关的关键所在 [5, 9, 22]。在这些论述中，何者可知、因而可计为风险，始终处于变化之中——它既受不断演进的测量技术、概率主义、意义建构和数据分析影响，也受到不同利益相关者之间争论与协商的塑造。

信用评分为考察算法系统中风险的建构提供了一个重要且影响深远的切入点：一方面，该领域的实践正在经历快速而显著的转型（尤其是在肯尼亚等全球多数世界地区，这些地方正如以下叙述所示，已成为新的金融与算法安排的某种实验室或试验场）；另一方面，这些实践至关重要，会对人生机遇、参与本地与全球经济的能力

（或无此能力），以及获取从银行业到教育、医疗保健等关键资源的机会产生真实而直接的影响。尽管信用风险的漫长历史跨越多种文化、国家情境和时间跨度 [44], ，算法信用评分的发展却深受美国历史与监管变革的影响 [22, 63, 93]。[1] 19 世纪末兴起的信用评分，建立在利用个人和商业信息，根据借款人的道德品格和财务状况对其进行识别和分类的基础之上 [63]。然而，直到 20 世纪 50 年代，“个人信息得到充分标准化、计算机化，并被转化为客观数据”之后 [57, 第 33] ，信用评分系统才“将信用worthiness重新定义为抽象统计风险的函数” [63, 第 183]。Fair Isaac Corporation 推出的 FICO 成为行业标准，将费力收集的个人信息纸面记录转化为代表信任和信用worthiness的三位数统计指标。信用评分模型不仅改变了放贷决策，也催生了新的商业模式，包括保险、汽车贷款和抵押贷款中的基于风险的定价 [57]。据 Lauer 所述 [63], ，这些发展引发了关于消费者保护、数据偏见和数据隐私的重大争论，并促成了具有范式意义的监管干预措施，例如《公平信用报告法》（1970 年），该法限制有关借款人的信息，并要求对拒绝信贷作出透明说明；以及《平等信贷机会法》（1974 年），该法规定不得基于种族、性别、国籍或宗教等受保护属性拒绝提供信贷，并要求对信用评分决策作出解释。随着 20 世纪 90 年代数据挖掘的发展，《格雷姆-里奇-比利雷法案》（1999 年）针对金融信息在公司内部及公司之间的流转设立了新的透明度规则，其中包括披露在一种情境下收集的消费者数据被共享并用于另一种情境的相关要求（这一问题后来被 Nissenbaum 理论化为情境完整性 [87]）。

尽管算法信用评分在不同行业和国家经历了多个转译、稳定化和标准化阶段，但自新千年之交以来，随着技术变革、经济与政治动荡，以及基于风险的评分日益应用于生活的各个方面，新的问题和担忧也随之出现 [41]。例如，9·11事件和2008年全球金融危机之后的危机模糊了风险与不确定性之间的界限，其影响显而易见地体现在各种预测系统的设计中。不确定性问题已变得与可计算的风险类别同等重要：“被伪装成风险的不确定性依然是不确定的” [22]。与此同时，这些危机催生了新的跨国数据与技术治理模式，这些模式源于国家与非国家行为体之间多元的制度安排 [36, 47, 113]。政治地理学家 Louise Amoore [5] 对这一转变作出了引人入胜的阐述。她指出，风险虽是连接当代安全与经济领域的一项主要技术，但风险的表述方式随着持续的经济危机和安全需求而变化，也呈现出越来越接近不确定性的特征。风险具有投机性，其运作超越概率范畴，进入（不确定的）可能性领域 [5, 109]。在这种数据治理模式下，人人都被界定为风险主体，但风险的构建不再围绕“个人”，而是围绕“一个分裂主体，其风险要素甚至会在她自身内部将她分割” [5, 第 8]页。因此，制度化的数据全面搜集旨在捕捉风险的各个分裂面向，利用多种形式的数据，以及“直觉、情感、审美、道德和推测等非科学知识形式” [5, 第 9]页。社会学家 Fourcade 和 Healy [41, 第 83]页将此称为“数据命令”：对经济效率的追求被广泛数据收集项目的常态化所取代，“即使当时 [组织]尚不知道该如何处理已收集的记录”。这些项目极大地拓展了信用风险相关数据的范围，因为所有数据都被强行转化并重新用作信用数据 [3] ——借款人的生活世界也在数据的“映像”中被（重新）塑造 [17]。这一系列工作的最终结果，是由各类行动者群体设计并管理的一种新“可能性政治”：风险咨询公司的解决方案、跨国数据公司的服务，以及国际发展机构的干预等，都是对技术、经济和政治变革的回应 [5]。

正如人机交互、计算机支持的协同工作和人工智能伦理领域的研究所强调的那样， [26, 37, 39],这些变化反过来也体现在并落实于那些通常被置于背景中的工作形式及更广泛的背景之中，而问题正是在这些背景下被构建为供探究的问题。Fujimura [43] 指出，问题是否可行，取决于问题如何被界定，以及它们在多大程度上能够在不同利益相关者群体之间实现对齐。对齐只有通过连接三个组织层面才能实现：实验、实验室和社会世界。人机交互和人工智能伦理学者已将这一框架应用于数据科学和机器学习, [37, 39], 展示了数据和计算如何与分析方法（实验层面）、软件工具和平台（实验室层面），以及更广泛的制度与资金安排和法律结构（社会世界层面）相互作用。因此，对齐是支持不同实践层面之间协调与协作的技术性和论述性工作，这些实践由共同的问题和利益联结在一起（参见人工智能安全论述中的“对齐” [2, 51], ，其含义是约束模型遵循预先设定的人类目标和价值观）。

在信贷评分作为问题的情境中，对齐在一定程度上受到分析方法与工具进步的塑造 [4, 71, 79]；这些方法与工具聚焦于解释模型如何运作的解释性说明，进而影响更广泛社会世界中的各种组织实践与规范性考量。近年来，人们协同努力，试图通过一系列以人为中心的可解释性工具[11, 13, 65]来理解模型结构及其结果。这些工具促进了不同模型及其结果之间的比较，为技术设计提供正当化依据，同时也满足法律要求。例如，Fair Isaac 声称采用更简单的机器学习回归模型，因为这些模型更易于解释，且性能可与更复杂但不透明的模型相媲美 [40]。在对企业数据科学团队开展的民族志研究中，Passi 和 Jackson [92] 考察了“特殊融资”案例：信用评分受损或信用记录有限的借款人与美国的贷款机构进行匹配。在这一过程中，数据科学团队将一个“匹配问题”转变为一个

[1]在美国以外，中国 [66] 和英国 [1, 62] 等国家已根据各自的经济环境和政治格局，实施了定制化的信用评分版本。与此同时，《通用数据保护条例》和欧盟《人工智能法案》对人工智能驱动的信用评分施加了严格的法规，强调数据保护、公平性和透明度，以保障消费者权利并防止歧视。

“分类任务”，其中信用评分高于某一阈值的借款人（信用评分：>500）会被选中获得融资。尽管准确率的预测性衡量“在不确定性面前提供了计算上的宽慰”，[d] 但正当化与可解释性工作最终是协作性的，通过有组织的信任与怀疑、测量与评估以及可信度与怀疑实践展开 [92]。尽管对可解释性与性能改进的共同追求，是跨不同组织与社会世界实现对齐的主要方式之一，但人们较少关注这些世界以及其外部世界如何被重塑，以适应模型假设。

当预测系统走出实验室、面对现实世界时，会出现更复杂的不确定性形式。建模不确定性源于模型声称能够做到的事情与部署后所面临现实之间的差距。Kou 等人 [61] 指出，机器学习实践中的技术主张——例如经济效率、预测准确率等——往往脱离现实场景，因为机器学习会针对行业标准指标（例如基尼系数）进行优化，而这些指标对现实的呈现十分狭隘 [70]。作者提醒，不应假定只要纳入“足够多的社会层面输入”，并将其收录于具有社会和文化相关性的数据集中，就能轻易弥合这一差距 [61], ；他们认为，这种假设只会助长数据科学和软件工作中普遍存在的“改进文化” [14]。

情境性不确定性是对模型预测行使自由裁量权的产物。Pääkkönen等人 [95] 将算法系统比作官僚机构，并指出，尽管人们声称算法系统具有僵化性和理性，但它们实际上是自由裁量的空间，无法充分预见现实世界中的经验偶发因素。尽管这些系统有可能取代或增强人类能力，但它们并非以机器式理性的方式做到这一点。因此，聚焦不确定性有助于我们“分析自由裁量权如何与正式规则共同发挥作用，从而产生决策” [95]。作者还补充说，将各种形式的不确定性转化为可计算的风险因素，并不会消除人类自由裁量权，“反而会将其重新分配到 *[其他]* 不确定性所在的位置，而这些位置可能不为人所察觉”（原文斜体）。这些隐而不见、未被察觉的空间催生了变通做法，使系统得以保持“正常运作”，但也可能促成被视为对抗性的策略性行动。通常被称为策略性行为或“操纵系统”的做法，是算法系统的构成性特征 [125], ；用户会借此重塑系统行为，乃至绩效成功的定义。

总而言之，风险并非自然产生、可供测量的属性，而是一种历史偶然性建构。尽管人机交互（HCI）及相关领域重新关注社会技术系统中风险的理论与实践局限，但这方面的大多数研究仍将“风险”视为预测优化问题。然而，近期政治、经济与技术变迁极大拓展了风险的内涵，也模糊了风险与不确定性之间的界限。因此，从风险与不确定性之间的递归关系中，至少可以辨识出三种形式的不确定性：认识论 不确定性源于对可知和可准确测量之事物的认识有限（参见 Knight [59] 的早期研究）；建模 不确定性反映了预测算法在捕捉复杂现实世界动态方面的局限 [61]；而情境 不确定性则揭示了预测系统如何受到人为自由裁量空间变化的扰动 [95]。尽管大多数人机交互和以人为中心的机器学习研究，已在不同应用领域中预测系统的开发与部署实践方面带来了重要见解，但这些研究忽视了对齐所涉及的专业且务实的工作，尤其是在风险与不确定性深度交织之处。如果说风险指的是用于评估信用worthiness的稳定类别，而不确定性指的是超出或动摇这些类别的因素，那么对齐就是通过持续的实践工作来实现稳定的过程；不过，对齐本身永远无法真正完成：它始终是暂时的。

3 肯尼亚信用评分的历史

肯尼亚和东非的信用评分历史与当代实践，受到三股核心力量的塑造：1895年至1963年长期的英国殖民统治；过去30年来由小额信贷和移动电话推动的商业转型；以及殖民时期之前和之后一直存在的信任与信贷文化和制度形式。与其他地区的模式相呼应——例如，美国长期存在的“划红线”做法 [100]——整个殖民时期，信贷可及性和信用worthiness评估主要沿着种族和族裔界限组织，体现了制度和文化层面的不平等 [33, 112]；正如 Donovan [33] 所指出，世界银行1961年的一份报告发现，英国银行倾向于向亚洲商人放贷，而牺牲了 [东] 非洲人的利益，这表明其中既存在社会问题，也存在信息问题：“由于英国银行家不与‘非洲民众’交往，彼此隔阂太深，无法准确评估 [非洲人] 的信用worthiness” [33, 第108]页。与此同时，地方政府和殖民当局敦促民众培养“储蓄观念” [33] ，并为书面证据奠定基础，以便据此建立和核实信用worthiness [20]。这种做法以牺牲持久存在的“草根”结构为代价，例如轮转储蓄与信贷协会；这些协会在社会经济关系中运用具有文化特性的承诺与义务形式 [112]。殖民后发展干预措施（如小额信贷）借鉴了这些非正式逻辑，试图模糊敢于承担风险者（以及“创业型”人士）与“处于风险之中”的人（服务不足的贫困群体）之间的界限，将穷人描绘成只要获得小额信贷，就有能力成为自给自足的小微企业家。 [101]。格莱珉银行的小额信贷模式由诺贝尔奖得主穆罕默德·尤努斯于20世纪70年代在孟加拉国创立并推广，开创了一种新型社会借贷：小额贷款以“互助小组”的形式发放，这类小组通常由约四名女性组成，她们不仅彼此担任担保人，还共同投身于小规模创业 [101, 112].[2]

随着移动电话的普及以及 M-Pesa 的问世，肯尼亚的信贷格局经历了一个重要的转型阶段——

[2]在肯尼亚等地复制小额信贷模式的地方非政府组织和福利组织，擅长区分“好女人”和“高风险男性” [101]„ 确保贷款偿还不仅通过同辈压力等性别化手段得到强制执行，还借助地方酋长和警察的力量 [112]。然而，小额信贷的全球化（乃至金融化）也付出了沉重代价，侵蚀了其承诺的大部分社会影响和财务责任。2010年代，印度安得拉邦等地爆发的小额信贷债务危机，将信贷与绝望、高违约率以及自杀或死亡联系在一起 [101, 112]。

点对点汇款技术于 2007 [45, 83]由 Safaricom 首次在肯尼亚成功推出——这从根本上改变了消费者小额贷款的结构，使发展重心从小额信贷转移，并在“金融普惠”这一以发展为导向的旗号下，为新的实验性数字借贷模式开辟了道路。因此，肯尼亚当代算法信用评分的发展历程可分为三个阶段：i）M-Pesa 的成功，为数字信贷产品（如 M-Shwari）的运行提供了数据基础设施，同时银行和非银行机构内部的信用信息共享法规逐步制度化；ii）基于应用程序的信贷提供商转向使用替代数据（和机器学习技术），以摆脱 Safaricom 对数据基础设施的垄断控制，争取自由与自主权；iii）消费者数据保护法的编纂，以及“Hustler Fund”的创立——这项政府数字借贷计划于 2022年11月启动。

在第一阶段，Safaricom 与非洲商业银行（CBA，现为 NCBA）于 2012 年推出 M-Shwari，标志着一种新型储蓄与贷款产品的问世。该产品利用了首批有详尽文档记录的实验性算法信用评分模型之一 [25]。Safaricom用户可通过该产品获得由 CBA 提供支持、无需抵押的短期数字贷款，金额为 100 至 50,000 肯尼亚先令（1 至 500 美元）。M-Shwari 的使用量迅速攀升，但其真正的创新在于支撑该产品的广泛数据基础设施：在开发产品评分卡时，系统结合了通话详单（CDR）和话费消费模式。这些数据还与综合人口登记系统（IPRS）的数据相辅相成。IPRS 是政府建立的全国自动化数据库，存储公民的人口统计与生物识别数据；此外还包括 2013 [19, 34] 之后强制要求进行的 SIM卡登记所产生的数据。与此同时，肯尼亚中央银行（CBK）建立了信用信息共享框架，要求银行共享信用信息，最终促成三家征信机构（CRBs）获得许可：Tran-sUnion（2010年）、Metropol（2011年）和 Creditinfo（2015年） [69]。这些征信机构很快积累了大量负面记录（违约报告等），其中也包括自愿共享信息的非银行机构所提供的记录 [19]。到2020年，超过260万人在征信机构留下了负面记录，其中约15%的人所欠债务低于 2 [126] 美元。

第二阶段的特点是机器学习技术和替代数据得到广泛应用，Tala（2014 年）和 Branch（2015 年）这两家总部位于硅谷的初创企业相继入场，地区内其他非银行和非电信初创企业也纷纷涌现。这些数字贷款机构开发了由专有信用评分模型驱动、基于应用程序的数字信贷产品；这些产品不依赖封闭的银行或电信数据基础设施及其相关监管要求 [119]。尽管这些新入场者为部分人带来了获得信贷的新可能，但也引发了对消费者隐私泄露、激进债务催收和掠夺性放贷行为的重大担忧 [81]。正如数字借贷研究者所总结的那样 [35]：“[这些数字信贷应用]会温柔地把钱借给你，然后再来掐住你的脖子。”

在第三阶段，《肯尼亚数据保护法》（2019年）颁布，同时出台了《数据保护（一般）法规》（2021年），其中规定了隐私设计 和默认隐私 原则（尽管后续研究发现，在如何与第三方共享消费者数据方面仍存在持续差距 [82]）。该法还设立了数据保护专员办公室，以接收公众的投诉；与此相关，中央银行持续就一系列法规征求公众意见，体现了旨在使数字借贷正当化并加以保障的努力 [119]。与此同时，2022年，新当选的鲁托政府宣布推出其“Hustler Fund”，该倡议旨在向服务不足的青年以及微型、小型和中型企业（即所谓的“奋斗者”）提供数字贷款 [117]。尽管该基金由银行与电信公司组成的联合体支持，并广泛宣传其对信用评分的使用，但它也录得全国最高的违约率，不良贷款（NPL）率超过68% [55]。

尽管上述各阶段展现了不同的信用评分逻辑，但它们共同关注的是如何利用数据界定特定借款人及群体所面临的风险。这一微妙转变将个体借款人从其更广泛的社交网络中抽离出来，同时规定了不合规时的惩戒措施，从受到制裁的负面记录，到债务追收中的威胁与骚扰，不一而足。然而，正如 Shipton [112] 所指出的，围绕财富、信贷和义务形成的某些文化形态与地方实践，历经政治和技术变迁仍然延续，成为算法系统必须适应、却无法完全预见的背景。下文将结合大型数据集，以及对风险更广泛的定义和评估，介绍数据科学和产品团队在应对这些特定情境下的文化实践时所做的具体努力。

4 研究场地与方法

本文基于在肯尼亚内罗毕的人工智能与算法信用评分领域开展的为期九个月的民族志田野调查。2025年二月至2025年11月，第一作者（来自乌干达）居住在内罗毕，考察了四家初创企业以及参与设计替代性信用评分系统的各类金融和政策行动者的社会技术实践与机构实践。内罗毕被选为主要“田野场地”，一方面是因为它在金融创新方面取得了全球瞩目的成功，另一方面是因为第一作者熟悉当地的科技与初创企业生态系统；十多年来，作为一名科技记者，他一直关注并参与其中。尽管这项研究由第一作者主导并实施，对话者根据他的职业和学术背景，知道他是一位研究人工智能与信用评分的社会科学家，但他也得到了一组来自北美的合著者的支持与指导。这些研究者在非洲多个地区开展技术与发展项目研究和协作方面拥有丰富经验。

参与式观察在初创公司和银行的办公室、科技活动及政策研讨会中开展，也通过包括 LinkedIn、WhatsApp 和 Twitter/X [30]在内的线上渠道进行。2025年二月，Moonshot[3] 这家初创公司的高级研究员通过 LinkedIn 联系了第一作者，并对他此前关于东非移动货币的研究表现出浓厚兴趣 [28, 83]。他解释说，以人为本的洞见对 Moonshot 至关重要。这是一家以远程办公为主的公司，为受监管和未受监管的贷款机构开发专有的算法风险与信用评估产品，同时还运营一款基于应用程序的金融市场平台，将个人与数字贷款机构连接起来，以便获取贷款（同时也

[3]Moonshot以及本文中使用的所有初创企业、产品和参与者姓名均为化名。

领域	参与者 ID
人工智能与数据	P8 [M],P9 [M], P10 [M]**, P11 [W], P12 [M], P16 [W], P17 [M], P20 [M], P22 [W], P26 [M], P27 [W], P29 [M], P30 [M]**
金融服务	P1 [M], P2 [W], P3 [M], P4 [M], P13 [M], P18 [M], P19 [M], P23 [M]
咨询	P21 [M], P24 [M]
政策与监管	P6 [M],P14 [W],P15 [W], P25 [W]
发展与投资	P5 [W], P7 [M], P28 [M]

注： [男] = 男性； [女] = 女性； [ ]** = 两次访谈**表 1：来自五个主要领域的访谈参与者概况。样本包括 21 名男性和 9 名女性（N=30）**

从该应用程序获取个性化的财务健康建议）。观察的其他初创企业都是以远程办公为主、总部位于肯尼亚的公司，开发相关但各有特色的人工智能驱动信用评分产品：Startup A 将贷款管理系统与算法信用评分解决方案打包提供给非正规金融机构（例如在肯尼亚十分普及的储蓄与信贷合作社）；Startup B 提供技术系统，以克服设备端信用评分等任务中的计算和推理限制；Startup C 则为信用评分等解决方案开发借贷数据基础设施。Moonshot 规模最大、知名度最高，团队有 10 至 20 人，而其他公司的员工都不到 10 人。大多数团队远程办公，偶尔前往内罗毕的不同联合办公空间，或实地拜访客户，包括银行和非银行数字贷款机构。

通过知情同意，第一作者获得了部分准入许可，得以在初创企业的冲刺工作和公司活动期间进行现场参与式观察，也能前往初创企业正在建立合作关系的银行观察。由于初创团队及工作场所具有分散化和虚拟化的特点，这些观察零散且往往非正式，通常在初创团队来到城里时进行（每隔一个月一次）。我们关注的是构建这些人工智能系统的人们如何协作、如何共同思考工作，包括他们的数据收集和风险制定实践，以及这些模型如何依据多项技术基准和商业目标接受评估。在说明我们的研究兴趣后，我们通过线上和线下渠道与其他对话者展开交流，包括数据科学家、商业人士和政策分析师；他们在这些渠道中就信用评分及相关议题进行社会和学术讨论。鉴于内罗毕（以及其他科技中心）的初创企业往往流动性强、以网络为基础，这些方法延续了“社会环境”或“场景”分析的传统；从事高科技工作的民族志学者 Saxenian [107]、Neff [86] 和 Shestakofsky [111]都曾采用这种方法。虽然我们观察到的世界和实践可能体现了硅谷式的“数字创业”想象，但若认定对话者的工作和实践不过是简单复制或模仿，那就错了：本地行动者创造并改造了工具和实践，此后这些工具和实践已在全球传播 [42], ，而且至少在某些情况下，他们披露的实验性实践超出了监管较完善的信贷领域中已知的常见做法

这些实践在欧洲和北美的信贷市场中并不常见。事实上，一些在场的公司，包括总部位于欧洲或美国的公司，似乎被一种可能性所吸引：在一个限制较少、监管较宽松的空间开展实验，从中汲取经验并取得创新成果，再将其带回北美和欧洲——这种流动或方向性（“逆向ICTD”？）或许一直未受到人机交互和计算领域足够的重视（尽管人类学家 [24] 更早注意到了这一点）。在这些碎片化、分散化的信用评分实践领域，传闻与保密并存 [72]：例如，一些行动者想知道第一作者掌握了什么信息（想必是关于他们在信用评分市场上的竞争对手），或者他究竟想了解什么；与此同时，他们对自己的工作究竟涉及什么含糊其辞（以保密协议为由，对公司机密严加保守）。总的来说，多位参与者通过活动、访谈和在线通信，以正式或非正式的方式分享信息并相互交流——不过，一位潜在受访者拒绝了访谈请求。

第一作者记录并转录了超过 150 页的田野笔记和访谈记录，这些内容来自对上述初创公司及其他对话者的观测，还包括 11 场科技活动，以及包含 300 多个网页链接、屏幕截图和图片的档案。参加的活动和研讨会在主题和规模上各不相同，从小型本地活动，到肯尼亚顶级的年度初创企业与科技博览会，后者有来自肯尼亚、尼日利亚、乌干达、英国及其他地区的 3,000 多名参与者。第一作者还在共享办公空间及其他场合进行了许多非正式（未录音）的讨论和观测，并将其记录在定期撰写的田野笔记中，再与作者团队分享和讨论。此外，一家初创公司提供了一个匿名化数据集，其中包含超过 80,000 个关于肯尼亚信用评分的观测值（行）和 260 多个特征（列），以及一份私人报告和一份路演演示文稿。

此外，还开展了32次深度访谈（录音内容总计约38小时），其中8次以面对面方式进行，24次通过 Microsoft Teams、Zoom 或 Google Meet 进行，参与者共30人（具体分类见表 1）。虽然第一作者能流利使用英语和斯瓦希里语，部分对话和交流也以这两种语言进行，甚至跨越两种语言，但正式访谈均以英语进行。英语是肯尼亚的官方语言，在专业和非正式场合都很常用。访谈录音的文字转录

根据田野笔记和访谈整理的文本经过修订和编辑，以确保准确性和完整性，随后由第一作者在研究团队的协助下，依据人机交互领域扎根理论的原则进行分享和编码 [78]。通过对田野数据与田野中的关键时刻（包括技术活动和一次偶然的Twitter/X Space）进行分析性比较，研究采用主题编码，形成了“风险”“信贷文化”“叙事”和“金融科技从业者”等主要代码组，同时发展出贯穿本文的风险、不确定性和对齐等主题。以下各节呈现此次田野工作的研究发现，重点讨论替代数据的形成、风险的多重建构以及模型性能的协商。三个场景均以一则民族志田野笔记开篇（斜体），随后结合后续访谈和观测值，补充相关洞见与分析。

5 项发现

5.1 场景一：打造替代数据

这是一个[illegible]。从[illegible]雨就下个不停。在内罗毕一个高档地段的一间灯火通明、装有空调的会议厅里，正在举办一场关于金融开放数据的研讨会。厅内约有*50*位参与者，周围还有一个摄制组。研讨会现场的实时画面正通过主舞台上的巨型屏幕播放。活动正式拉开帷幕。主持人邀请组织方代表登上主舞台。来自行业组织联盟、审慎监管银行业的从业者、由肯尼亚中央银行（*CBK*）重新修订法规的一组非吸收存款的 *[数字]* 贷款机构，以及一个专注于发展、能够影响金融政策的机构的代表，与主持人一同登台，排成一列，彼此挽臂。主持人表示，这些机构希望携手合作，共享并利用彼此的数据；他还提到，上一次在肯尼亚举办此类研讨会还是在*2018*年。如今已是*2025*年，银行、电信公司、金融科技公司和信用征信机构（*CRB*）希望再试一次。随后，一位引导员受邀登台，并宣布当天的议程：金融开放数据。

随着幻灯片演示在巨幕上持续播放，主持人概述了关键议题。“开放数据以消费者为中心，”他宣称，“这是一种由消费者主导的数据关系。”他展示了世界各地开放数据模式的成功案例，首先以巴西为例：该国有*6800*万个活跃账户，其所生成的数据每周通过*“20*亿*+ API*调用”被访问。他以一个问题收尾，询问参与者：“开放数据要在肯尼亚取得成功，需要具备什么条件？”观众纷纷回应，一个人接着一个人，全都点名*Safaricom*。[4] 他们说，*Safaricom*已经演变成全国性的支付数据瓶颈，金融领域的开放数据只有在*Safaricom*的垄断终结之后才能成功。一名*Safaricom*代表为公司辩护，认为*Safaricom*促进了金融科技领域以及更广泛银行业的竞争。有几个人点头表示赞同，但连珠炮般的提问仍在继续。主持人插话建议，也许不该再谈*Safaricom*了。观众哄堂大笑。在下一环节，一位主持人提醒观众，并非所有数据都一样，并列出开放数据的三类数据：了解你的客户（*KYC*）、客户金融数据和市场洞察。尽管征信机构被授权收集并共享信贷信息数据，质量和覆盖范围方面的问题依然比比皆是。来自电信公司和银行的代表激烈反对共享“他们的”数据——他们称自己投入了大量资源获取这些数据，因此不会在拟议的开放数据框架下进行共享，尽管该框架强调消费者隐私。研讨会结束时，似乎还有很多话没有说完；为我争取到研讨会邀请的那家初创公司的一个人反思道，这次研讨会“泛泛”而“普通”。与我共进一餐的另一家初创公司的一名代表则夸耀说，他们积累的微型、小型和中型企业数据比征信机构还多。

这份田野笔记描绘了肯尼亚金融数据格局十多年来陷入僵局的状况：业界共享数据的努力始终未能形成足够势头，无法让所有参与者受益。征信机构虽已建立起规模庞大、所有人均可付费访问的数据集，但 Safaricom 对远为优越的实时数据基础设施的垄断却具有排他性。然而，新兴企业正设法通过其他途径获取 Safaricom 的数据（以及其他来源的数据），而且似乎已有进展。下文将展示，一群行动者如何运用技术手段收集并分析海量数据，同时也将指出替代数据基础设施的脆弱性。

5.1.1 数据收集与分类。 我们访谈的多家初创企业表示，他们正在采用技术变通方法，实时收集金融交易数据，尤其是针对被视为“信贷记录薄弱”或“信用不可见”的人群（其信用信息未被征信机构记录；其中大多数属于青年和农村妇女等群体）。通过这些方法，数据收集可以自动化，直接从用户设备抓取数据，或通过与第三方数据平台相连的应用程序接口获取数据。在这些方法中，研究的许多参与者都认为“短信抓取”是黄金标准。访问用户的短信/文本消息，可以获取大量有关手机用户的交易与行为信息，涵盖公用事业、食品和医疗等领域的交易。Startup C 首席执行官 P11 表示，短信日志是重要且易于获取、可验证的个人数据来源：

> “所以我们问自己，最丰富的金融记录在哪里？在新兴市场，最丰富的金融记录其实就在短信里，*[*比如肯尼亚*]*。好吧，在美国，最丰富的金融记录则是电子邮件。每次你在美国买东西，都会收到一封确认邮件。所以，如果我想在美国构建同样的基础设施，就会开发一套基于电子邮件解析的机器学习基础设施，因为那里存有最丰富的金融记录，你支付的所有款项都会记录在其中……新兴市场以及肯尼亚这类地方的优势在于，我们有移动货币，*[*例如*M-Pesa]*，，它由电信公司提供服务，并掌握着大量数据。”– [P11], Startup C 首席执行官

初创企业通过抓取并解析短信数据，挖掘 M-Pesa 的消息提醒（交易后发送给用户的提醒），实时重建 M-Pesa 对账单，

[4]尽管 Safaricom 是肯尼亚领先的电信公司（在全国 5600 万人口中拥有 5000 万名用户），但它处于法律监管的灰色地带。它并非受审慎监管的金融实体，然而 M-Pesa 及其一系列金融产品却占据了数字信贷市场 95% 的份额。除了金融科技之外，Safaricom 也是肯尼亚安全与官僚体系的核心，并深刻影响着人们对进步与创新的想象。

以及银行和其他服务提供商发送的提醒。虽然通过这些变通做法获取的数据有时并不完整，但与其他数据结合后，已足以生成有用的洞见。

> “通过再次与金融科技公司合作，银行也取得了一些进展，尝试了解如何收集这些信息。我相信你可能也用过一些能够通过软件开发工具包（*SDK*）*[窥探手机数据的应用]*。当然，你必须在应用中正确设置权限，才能允许它们窥探数据。因此，银行与金融科技公司合作，开发了一些应用，能够从你的*SIM*卡中窥探数据，查看你的*M-Pesa*账单、话费充值、其他银行提醒。我们试图重建*Safaricom*掌握的同类信息，但当然，这并非百分之百准确。不过，这也让银行有机会分一杯羹。”– [P3], 某银行的产品负责人，拥有数十年经验 pe- 银行业经验

开发短信解析器以进行数据收集所需的技术专业知识，反过来又依赖于文化和情境性知识，用以解读、标注和分类所收集的数据。例如，尽管Moonshot已使用其软件开发工具包套件从用户设备收集了超过5亿个数据点，由首席技术官和初级数据科学家组成的团队仍投入时间来解读和分类所收集的数据。P10说，他曾在共享办公空间的一个角落里花了四个小时，手动阅读与财务报表交易相关的数据，并试图弄清其中究竟发生了什么：“如何从中提取这些信息？ *[...]* 人们如何花钱？人们如何赚钱？他们会博彩吗？他们会储蓄吗？”。一家公司委托另一家公司来支持标注和分类方面的解读工作，也并不罕见。例如，P8向一家资金雄厚、希望打入肯尼亚数字信贷市场的硅谷初创公司提供数据标注服务。P8的团队获得了一个包含50,000条短信、采用可读文本格式的数据集，并收到明确指示，要求识别特定实体和类别。除“时间”、“名字”、“姓氏”、“金额”和“公用事业”等通用实体外， [P8] 还凭借当地经验建议增加其他实体，包括“体育博彩”和“食品”。

数据管道建成（并实现自动化）后，对话者表示，他们可以通过多种方式收集和分析数据。客户数据档案可随时调取，也可以生成综合报告，“实时生成，只需*20*分钟”，P11补充道：“客户数据档案无所不包：他们的偿付能力、所贷的款、是否赌博、在哪里吃饭 *[.....]* 对我们来说，信贷就是你必须做的事，*[为了维持生计]*。这些数据能发挥的作用太多了，对吧？你知道，甚至还能涉及医疗保健。”

5.1.2 监管要求与技术依赖。 我们的对话者谈到了促使其数据来源多元化的行业压力，包括通过与政府机构（例如农业数据或卫星数据）乃至体育博彩公司等行动者建立正式和非正式的数据共享合作伙伴关系（这些已积累了极为可观的交易数据量）。世界银行和国际金融公司委托撰写的一份关于肯尼亚“替代数据”的未发表报告，截至2025年7月以草案形式流传，着重指出信贷获取不足是消费以及微型、小型和中型企业（微型、小型和中型企业）增长的主要障碍。报告显示，仅有40%的微型、小型和中型企业被纳入传统征信机构数据；若使用替代数据，这一比例可大幅提升至75%。报告指出，通过“消费者主导、基于同意的数据共享”，机器学习的新发展正在使“这些数据能够以相对较低的成本被提取、转换并安全地与第三方共享”。由于信用信息共享方面的法律对替代数据的合法性规定并不十分明确，对话者认为，这些收集额外数据的努力对高度监管行业（例如征信机构，世界银行曾建议其加大对替代数据的投入）施加了压力。但仍有回旋余地（甚至增长空间）。据肯尼亚一家数据交换平台的首席执行官P12称，需要更多样化的数据来源，包括“实时数字支付”和“Hustler Fund贷款详情”。

> “征信局在追踪已办理的 *[债务]*产品还款方面很有用，但它并非扩大覆盖范围或提升信贷获取的全部答案，因为实际上还有很多信息缺失，未被征信局记录 *[.....]* 因此， *[征信机构]* 有很大的机会，能够让银行的资产负债表扩大两三倍；具体幅度取决于这些替代数据。由于可见的信息有限，资产负债表相当薄弱，因此存在很大的增长机会。”—— [首12], 数据交换平台首席执行官

然而，这些多样化的数据来源又依赖于其他静态数据来源，包括个人信息（出生日期、性别、国籍）和金融数据（SIM卡号码和银行账户）。静态数据来源提供了急需的了解你的客户信息，可用于验证个人身份或金融身份。一位受访者指出，部分一级银行依赖Safaricom的碎片化基础设施。曾在一家大型电信公司工作过的P19表示：*“[在] Safaricom [...]* 中，没有一个单一的数据来源能够真实反映客户数据，或我们所称的主数据。” 这意味着数据以一种不太结构化的方式从多个内部数据库中提取。这些碎片化基础设施在后续环节引发了更多问题：由于SIM卡号码具有可转让性，金融身份变得不可靠（某个号码在一段明确的停用期后可以转售给另一个人）。同样曾在一家大型电信公司担任工程平台经理的P23表示：

> “大多数信贷机构都以 *[SIM*卡*]* 电话号码作为 *[*获取和验证借款人身份以进行放贷*]*的依据。如果一个电话号码 *[.....]* 曾对应过*6—7*个身份，而其中*80%*的身份都使用过*[*数字贷款*]* 服务，那么有时它就会成为一个有缺陷的 *[*数据*]*衡量参数，因为生命周期管理，尤其是在贷款机构一侧，往往没有得到妥善管理。”

本节有三点启示。首先，对替代数据的寻找几乎没有止境，需要运用各种技术和文化专长。任何可以收集并分类的记录

都会被收集和处理。其次，替代数据极大地扩展了记录的定义，尤其是对于那些可能被传统数据基础设施忽视的人群而言。正是通过这种扩展性的审视，替代数据得以实现不同形式的区分。第三，替代数据可能在接缝处断裂。它们可能十分脆弱，因为新出现的或长期存在的技术、法律和制度问题会削弱其有效性。

5.2 场景二 – 剖析风险

本周即将结束。在内罗毕一处高档活动场地中，一座敞开屋顶的场馆角落里有一个小舞台，四周放置着四台冷焰火机，白色火花从中喷涌而出。场地里挤满了中年男性，另有少数女性和年轻人。大多数人穿着商务休闲装，不过也有少数男性身着西装、系着领带。有些面孔（对我而言）颇为眼熟，因为我曾在城里不同的科技活动中见过其中一些人。三位创新者笑容满面地站在舞台上，身后彩纸炸开，一首新近发布的非洲节拍歌曲在空中萦绕。这是重要的一天——一场仅限受邀者参加的活动，面向数字借贷领域的人士（尤其是银行、电信公司和资产融资机构），他们齐聚一堂，见证一项新信用评分的发布：经过数月的传闻与期待，*Extenscore*® 终于落地。该评分由 *Moonshot* 与肯尼亚一家成熟的数据交换平台合作开发。产品负责人宣布，这项协作已持续一年多，双方围绕研究与实验投入了大量精力，以确保评分准确无误。负责人补充说，他认为这项新评分将如何支持“社区”，与他此前在海外银行业的经历密切相关。他说，他曾任职的银行通过利用教会等社区关系，以安全的方式向“吉普赛人”和“非法移民”以及没有固定住址的人发放信贷，从而生成关于这些人的数据轨迹。他惊讶于“信贷的重要性”竟能在多大程度上改变人们的生活。他说，像数据这样无形的东西竟能产生具体成果，使人们得以获得“设备”和“货车”，这具有变革意义。随着讲话继续，热情与溢美之词不断升级：革命性的、创新的、破纪录的。他邀请首席技术官上台，并连同团队一起称赞其为“非常懂技术”。

首席技术官在热烈的掌声中走上舞台。他解释说，这项新评分结合了领先信用征信机构提供的金融数据，以及 *Moonshot* 从借款人的手机元数据和短信日志等来源获取的替代数据。他宣布，该模型已基于一个包含超过 *1000* 万笔匿名化交易的数据集进行训练，性能远超三种行业标准评分，优势达数个数量级。这些对比评分并未被点名，但所有人都能看到相关统计数据，它们显示在小舞台右侧的大型像素屏幕上。首席技术官对这些数据作出解读。他说，在区分优质与不良借款人方面，该评分的表现比竞争对手好 *80%*。这项评分也具有包容性，能够服务于微型、小型和中型企业、青年群体以及农村地区的人们；长期以来，这些群体在征信机构的传统金融数据中一直代表性不足。后来，庆祝活动渐渐转为交流联谊和非正式交谈：“要是我当信贷专员时就有这项评分就好了。”“我曾在伦敦打造过类似的决策引擎。”“如今的机器学习模型变得更容易解释了。”无论是陌生人还是朋友，都在分享这些故事；随着首席技术官在人群中穿行，故事也不断流传，他邀请来宾们彼此交流，并期待 *Moonshot* 未来推出的新产品。

在这份实地笔记中，一种新的信用风险评分高调亮相。根据其开发者的说法，正是凭借他们对风险的精通，以及数据科学和机器学习的高阶技术，他们才取得了如此令人印象深刻的表现结果——至少在纸面上如此。通过幽灵般的叙事与性能表现，并借助与更成熟的行业行动者建立的合作关系，为风险主体革新数字借贷的承诺得以在话语层面实现，并获得正当性。在以下小节中，我们将展示“风险”是如何被建构的，包括为支持其合法性及持续运作而必须不断检验的各种假设与假说。

5.2.1 专家猜测如何预示特征工程。数据是构建信用评分模型的基础。正如我们上文所展示的，收集和分析数据需要投入大量工作。然而，这些数据所支持的模型假设，反过来又依赖于高度结构化的数据集，用于提取特征（变量的属性）和收集标签（关于过去贷款表现的决策）。虽然特征提取和特征工程属于技术性成果，但标签的初始数据往往从信用参考局获取，而其数据覆盖范围较差。因此，大多数行动者都不得不从头开始，尽管他们获取资源的程度各不相同。我们的对话者回顾开发信用模型的初始阶段时，将这一过程描述为“假设检验”和“专家猜测”。正如一位参与者所说的那样，他在信用评分的最初阶段曾就职于一家大型银行，如今则在非洲多个市场担任信用评分顾问：

> “你需要示例数据 *[*来构建模型*]*，而在*]*起步阶段 *[*的主要问题是：你尚未推出服务，因此没有示例数据。那么该如何起步？你必须先提出一些假设。在信用评分中，有一种方法叫作专家评分卡。你可以根据假设制定专家评分卡，并将这些假设视为未经验证的猜测 *[.....]* 说实话，这就是猜测，不过是专家的猜测。”– [P30],数据科学家和顾问

专家评分卡过去在银行中更为常见（不过，初创企业在起步阶段也会使用这类评分卡）。经验丰富、任职时间较长的信贷专员和风险分析师会制定评分标准，为独特的文化和行为特征赋予任意分数。

只有在积累了有关贷款表现的数据之后，特征工程才使我们的行动者能够“拆分重组”[P30] 借款人行为变量，将其细分为多个子变量。四家初创企业的数据科学家和工程师向我们介绍了如何从包括收入流入与流出在内的诸多变量中提取特征，并将这些特征与当地及具有文化相关性的行为联系起来。尽管根据个人偏好可以构造出数百个特征（例如，按时间对收入进行细分：“上个月的收入、两个月前的收入、三个月前的收入，等等。”），但目标始终是将特征数量控制在可管理的范围内，通常少于 20 个：那些夸耀拥有数千个特征的公司其实是在进行“公关活动”，[P10]。然而，选择模型中应使用哪些特征的过程，实际上是检验特征之间相关性和强度的过程

这一过程被 [P9] 比作一套“酷炫技巧”，依赖于“不可扩展的人工检查”。特征工程的这一阶段既需要创造力，也需要统计严谨性，必须不断迭代，以确保捕捉到最具预测性的特征。

> “因此，在银行业，我们通常会使用带分箱的逻辑回归，方法很简单。我们会使用一些软件工具，主要是 *optbinning*。接着反复迭代，直到对特征 *[...],* 满意为止，然后拟合 *[*模型*]*。”– [数26], 一级银行的数据科学家

5.2.2 申请评分和行为评分如何依赖专家猜测。 几位参与者将信用评分模型分为两大类：申请评分模型和行为模型。申请评分模型用于评估（新）借款人偿还贷款的可能性，而行为模型则利用已经建立的信贷关系，预测现有客户偿还贷款的情况。尽管有人担忧信用参考局评分的数据质量、覆盖范围和完整性，但几家商业银行和新兴数字贷款机构仍高度依赖这些评分来生成申请评分卡。在我们的研究中，参与者花了大量时间分析模型中特征与标签之间的相关性，但这些模型只有在真实场景中部署后才真正接受检验。在一项由一家新兴数字贷款机构资助的早期实验中，该机构热衷于利用 Moonshot 的替代数据和机器学习评分技术；然而，Moonshot 对其初始申请评分模型的表现感到震惊。

> “我们做的第一个模型非常简单。它统计了过去六个月的资金流总额，以及资金流总额，还有资金流出的标准差之类的指标。还款率糟糕透了，真的很差，大概只有 *20%*。当时我们只是在做实验，所以这也不是什么大不了的事，但我们的反应基本就是：哦，天哪。实际上，我们当时觉得，现金流活动非常频繁、金额高达 *150* 万肯尼亚先令的人，肯定会还款，尤其是因为我们刚开始提供的小额贷款金额是 *KES 500]*。显然，收入越高，风险就越低。但实际上，这往往根本不是最主要的特征。高收入并不是最具预测力的特征。实际上，现金流活动频繁的人并不怎么在意你要罚他们的那些小额罚金，所以他们会想什么时候还就什么时候还。他们同时向五六家不同的贷款机构借款，然后会优先决定先还哪一家。这背后的心理因素实在太多了。”– [P10], Moonshot

Moonshot 遇到的挑战并不少见。其他参与者也分享说，他们同样认识到，信贷决策（批准或拒绝贷款）必须建立在更深入理解风险和行为的基础上。Moonshot 则更多借助用户调查和焦点小组来了解借款人心理，尤其是在申请（获客）和行为（留存）模型方面。

其中一项干预措施是彻底摒弃二元分类方法（会违约 *vs.* 不会违约），转而分配局部最优信贷额度。信贷额度采用浮动额度，根据每位借款人的风险画像量身定制。

> “因此，我们认为有两个人各自可以获得 *20,000*。他们的行为相似，等等，只是也许过去的一些 *[*使用模式*]* 告诉我们一些信息，让我们或许需要对此 *[*保持*]* 谨慎。因此，我们相信他们两人都有还款能力 *[*但*]* 我们觉得*[*其中一人*]* 可能不会还款。 *[*所以*]* 我们基本上会重新调整他们可能获得的最高额度 *[.....],* 我想从某种意义上说，这降低了贷款人的风险敞口，因为对于风险敞口更大、风险更高的人，我们会先给他们较低的额度，然后随着时间推移，如果我们发现他们的行为有所改善，他们就能重新获得 *20,000 [.....]* 如果我下调你的额度，并不是在惩罚你，只是在限制你所拥有的偿付能力。”– [P10] , Moonshot

然而，据我们的访谈对象所说，设定信贷额度也带来了新的“风险”。人们担心借款人会试图“钻系统的空子”，不断提高信贷额度，这种担忧被视为始终存在的“对抗性问题” [P10]。管理这一问题依赖于对欺诈行为的文化和行为假设，人们也试图防止此类行为。

> 所以，如果用户进入 *[*数字贷款*]* 应用并尝试申请 *[*贷款*],* ，而该设备在过去 *10* 天内被用于提交了 *20[*笔贷款*]* 申请，你知道，还用了四个不同的电话号码，这就很奇怪。所以，我们关注的是你在征信局报告之外的 *[*更广泛的*]* 数字足迹。” – [P10], Moonshot

> “我们发现，如果我们使用交易数据，而客户明白如果账户里有很多信贷、经常使用账户， *[*额度*]* 就会提高，那么这就容易被操纵。为了防止这种情况，他们在参数中引入了其他选项。其中有一个特征选择模型，即使你尽可能多次地钻空子，额度也不会超过某个百分比，因为还要考虑其他特征*[.....]* 我们使用 *85%* 的信用参考局评分和 *15%* 的内部评分来计算你的最终分数。我们还会评估该账户的偿付能力。如果这个人的偿付能力是 *1,500*，我们就可以放心地提供 *1,000* 的贷款。”

– [P1], Data scientist at Tier 1 bank

试图限制“对抗性行为”的努力与反馈循环，遭遇了数据和模型的根本现实：漂移。在数据科学和机器学习领域，由于用户行为发生变化，导致模型预测能力随之下降，数据和模型会随着时间推移而衰退。Moonshot 至少有 15 个模型，其中大多数是行为模型，服务于以下产品：“超短期贷款和 *7* 天期贷款、*28* 天期贷款、*3* 个月期、*6* 个月期、*1* 年期、*2* 年期贷款。还有汽车贷款、车辆登记簿贷款、资产融资， *[*等等。*]*” [P10]。一家一级银行至少有 25 个模型，服务于不同的产品细分领域：“我们有面向消费者的产品，也有面向中小企业的产品，包括中小企业个人客户和中小企业商业客户 *[.....]* 所以

此后每种产品都有自己的评分卡" [26]。我们的大多数参与者都会定期更新模型，以应对这些漂移。一家一级银行的从业者表示，他们每月都会更新所有贷款类别的行为模型目录 [P26]。Moonshot 更新模型的频率较低， [P10] 认为，与其定期更新模型，他们更注重借款人"心理因素"方面的基本假设；这让他们对自身的数据和模型更有信心，也降低了对持续更新的重视。

本节有两个关键见解。第一，风险是通过专家猜测和情境化的专业知识来界定的，这些知识会对数据进行"拆解重组"，从而形成对风险的多种解读。第二，从业者在设计时会预设借款人行为，导致普通行为被重新归类为可疑或具有对抗性，大幅扩大了对消费者进行监控的范围。

5.3 场景三——平衡度量指标

一位肯尼亚政府成员正在 *Twitter/X Space* 上发言。用户可以在这个平台免费主持类似电台谈话节目的公开语音节目。两位社交媒体名人正在采访中小微企业发展国务部常务秘书 *Susan Auma Mang'eni*。她正在阐释 *Hustler Fund* 的"概念内涵"，称该基金"是一项旨在解决信贷可及性市场失灵问题的政策干预"。在接下来的*34*分钟里，常务秘书持续发言，阐述该基金的宗旨，并强调它旨在弥补的市场缺口。她谈到肯尼亚的大多数无银行账户者，指出该国非正规部门人口近*1,700*万，而正式就业者仅有*350*万。她说，传统银行认为非正规部门从业者风险很高，因此他们往往无法获得金融服务。她形容这些"奋斗者"的经济处境如同生活在"抽奖局里：今天有点收获，明天就没有了"。常务秘书接着解释，该行业的许多人依赖数字贷款机构，而这些机构的日利率高达*10%*。"这是'数字奴役'！"她宣称。不过，她也指责肯尼亚人是"惯常违约者"，并指出违约者一旦被征信机构列入负面名单，其记录就会保留七年，进一步阻碍他们获取信贷。

她表示，*Hustler Fund* 旨在解决这些"市场失灵"问题。她指出，**254#* 是通过移动电话访问该基金的 *USSD*代码。她兴致勃勃地提醒听众，这也是肯尼亚的国家区号。这个仅提供数字服务的基金，可通过所有主要数字渠道免费使用，由政府提供资金，并由电信公司和银行组成的联合体负责管理。她点名提到这些电信公司：*Airtel*、*Telkom*，以及"甚至还有 *Safaricom*"。该基金利用联合体的数据为借款人开发信用评分。她建议说，借助"数据"，"你的行为就是你的财产"。她强调，这对于"扩大……基于风险的融资范围"至关重要。新借款人的最低信贷额度为肯尼亚先令 *500 [约合美元4]*，还款期限为 *14* 天。她指出，这个金额或许看起来很少，但她辩称，肯尼亚是一个"小微经济" *[小规模经济体], , 98%* 的商业都是微型、小型和中型企业。她补充说，这一比例与其他数字贷款机构的数据一致，该基金只是在顺应市场趋势。如今，该基金拥有全国最大的借款人数据库。尽管该基金目前有 *2600* 万借款人，但她关注的是其中 *500* 万名"还款表现非常好"的借款人。主持人打断这位常务秘书，请她用"更贴近人的方式"讲述这个故事时，她改变了语气："这一整周，我们都在收到这样的消息……请不要解散这个基金。"

上面的田野笔记记录了一次在 Twitter/X 上偶然发生的相遇：一场持续了两小时零七分钟的语音空间。PS 发言时间超过 90%，且常常未被打断，详细阐述了基金的理念，以及自2022年11月 启动以来取得的卓越成效。约有 75 人参加了这场语音空间，而它直到播出当天早些时候才做宣传。然而，那周早些时候，一个著名的民权组织发布了一份措辞尖锐的报告，抨击 Hustler Fund [55]，"辜负了奋斗者"。报告指出，违约率高达 68.3%，并以监管漏洞和政治妥协为由，呼吁立即"废除"该基金。接下来的故事将不再围绕语音空间中受政策启发的独白展开，而是展示参与者如何探讨度量指标在衡量数字借贷信用评分表现时的功能与作用，以及从业者裁量权得以发挥的场景。

5.3.1 失败何时是合理的。 据我们的对话者（包括几位直接参与者）所述，Hustler Fund 正是政治介入基于规则的算法信用评分过程时会发生什么的一个例子。正如肯尼亚一家领先电信公司的工程平台经理 P23 所解释的：

> *"M-Pesa* 是开发 *Hustler Fund* 的核心。因此*[M-Pesa]* 确实与合作银行合作，当然 *Airtel money* 也参与其中，还有 *Telkom* 的 *T-cash*，但最核心的是 *[M-Pesa], ,* 它就像国家支付交换系统 *[...]* 因此，向 *[M-Pesa]* 提出的挑战是，现有解决方案所做的信用评分将大多数人口排除在外 *[...]* 既有信贷机构 *[*贷出*]* 的金额非常低 *[...]* 有些情况下，某人仅符合获得 *100* 先令的资格。因此，这意味着 *[*新的*]*评分模型将偏离现有贷款机构的做法。就信用评分而言，赌博 *[*即体育博彩*]* 是一个重大议题。客户账户上的交易额也是一个重要项目。于是，所有这些在某种程度上都被搁置一旁。"

然而，基金的表现十分糟糕。P23 强调，尽管移动贷款的行业还款率约为 60%，Hustler Fund 的还款率却"远低于 40%"。造成这种情况的原因有很多，包括移除了体育博彩等重要特征。其中值得注意的是，违约时不再留下负面记录："违约者不会被 *[Hus- tler Fund]* 列入征信机构的记录。这实际上是总统亲自传达的。" 其他开发信用评分模型的行动者表示，是否纳入某些特征取决于开发者偏好。尽管体育博彩（赌博）被视为一种负面信号，但包括年轻人在内的大量人群都参与其中。认为体育博彩与拖欠相关并非毫无根据，因为最常参与体育博彩的人更有可能违约。不过，一些行动者表示，他们已经找到了解决办法： [P10] 暗示，Moonshot 追踪了另一项重要特征（出于专有原因，他们要求不要公开），其反向关系使他们能够

将体育博彩纳入考量（同时避免赌博相关的负面信号扭曲其结果）。

不过，P23 承认，尽管贷款偿还等参数会在借款人的行为评分中发挥作用，但从一开始，商业人士 [联合体中的] 人就将 Hustler "视为一项政治承诺，而不是一项商业决策，并且也确实按照这一点执行了。因此，对于这些企业来说，他们或许赚到了钱；但就整个 *Hustler Fund* 本身作为一个能够长期持续的商业案例而言，我不认为当时是这样考虑的。"

5.3.2 当技术评估遇到裁量。 在经常被公开引用的不良贷款率（这一度量指标）之下，还存在一整套技术性的数据科学和机器学习度量指标，其中包括基尼系数，用于评估正在开发和部署中的信用模型的表现。数据科学领域及与数据科学相关领域（尤其是产品领域）的参与者始终指出，基尼系数在这里被用作一种统计指标，用于衡量不同模型在预测违约方面的相对成功程度。[5] 基尼系数衡量模型区分还款（良好结果）与违约（不良结果）的预测能力，取值范围为 0 到 1；数值越高，表现越好。不同模型类别的基尼系数可接受范围有所不同：正如 P26 所解释的，" 一个良好的行为模型的基尼系数为 *0.6*，而一个尚可的申请评分模型的基尼系数至少为 *0.3*，因为数据较为稀缺。"

银行的产品和业务分析师对基尼系数指标与现实环境中的表现之间长期存在差距感到沮丧。一位曾在大型电信公司工作、如今已在银行业工作二十多年的产品负责人分享了他的担忧。

> "过去我一直很困惑：你会发现一张评分卡的基尼系数总是很高，比如 *80%* 或 *85% [0.8 - 0.85]*. 但到最终评估表现时，基于这些评分卡发放信贷后，你会发现违约率在 *[30%* 到 *] 15%* 之间。这是理想情况下的预期结果吗？可能不是。"– [P4]
>
> ，银行产品负责人

根据上述内容，按照基尼系数公式，[6] 所声称的80%–85%通常应对应3%–6%的违约率（而非10%–30%的违约率）。对于报告中约10%–30%的违约率，其基尼系数可能较低，从负值到最高25%（或0.2）不等。 [P3]和 [P4] 指出，数据局限长期困扰着数字借贷，并推测来自信用参考局和内部（银行）记录的信用数据并不充分。尽管增加数据被视为弥补信用评分不足的具体解决方案，P4认为，即使是Safaricom这样拥有海量数据的大型参与者，其申请评分模型也面临着与行业其他参与者相似的挑战。归根结底，提升贷款产品整体性能指标所依靠的，是那些"优质"客户。众所周知，扩大规模（用户、贷款账簿、数据等）不仅能增强信贷评分的预测能力，还有助于在已知市场现实（违约情况）的背景下重新平衡贷款组合。我们的受访者呼应了许多其他人的说法，并更具体地阐述了这一观点：

> "肯尼亚市场的大众市场申请评分卡会产生 *20%* 到 *30%* 的违约率。如果你放贷，*20%* 的人不会还款。好处是，那些会还款的 *80%* 之后通常也会继续还款 *[*而且*]*你会不断积累回头客，对吧？那么，你要解决的问题就是：你是否有一张评分卡，能将违约率从 *20%* 降到比如 *15%* 或 *10%*，因为你不可能把它降得更低，对吧？ *[...]*所以，你必须依靠优质客户来弥补不良客户造成的损失 *[...] [*像 *Safaricom* 这样的业内企业*]* 花了大量时间争取优质客户 *[*，并可以*]* 在不使用任何评分卡的情况下向他们放贷。他们都是优质*[*客户*]*，，但你仍然需要不断发掘新客户，对吧？这就是评分卡的价值所在。但问题是，代价有多大？"
>
> – [P4], 银行产品负责人

谈话往往不止于基尼评分，还会回到自由裁量的空间：商业和放贷决策是在既有社会规范、人际关系和义务的背景下作出的。由于管理和跟踪信用评分的文化尚未（或还未）成为日常信贷体验的一部分，解读并转化这些评分的工作便依赖于产品和商业经理对社会与文化的了解（例如，对一般贷款违约动态的理解）。据 P8 介绍，他此前曾在国外开发风险模型；肯尼亚与西方国家之间信贷文化的差异必须结合具体背景来理解：

> "我认为，*[*信用评分*]*根本毫无意义，因为从根本上说，它们无法准确衡量风险 *[.....]* 如果这是我第一次申请信贷，你知道，这并不意味着我是个糟糕的借款人，因为我没有任何信用记录。不过，你也得明白，在这里的文化背景下，我没有 *[*正式的*]* 信贷反而是件好事。"

有几个例子表明，金融领域的行动者会忽略信用评分，或只部分参考这些评分：在内罗毕一家投资基金担任分析师的 P7 表示，CRB信用报告中的信用评分"还可以"，但未必能说明某个投资机会为何不值得投资。人们会花更多功夫，从更个人的层面了解这位企业家。另一个例子中，P23 将这些个人化和关系导向的做法归咎于"我们如今看到的不良贷款率居高不下的一个重要原因 *[*今天*]*。"借款人（包括在信用参考局有负面记录的借款人）会利用自己与储蓄信贷合作社董事和信贷专员之间的关系来获得贷款服务，尤其是在教会或社区组织的储蓄信贷合作社中。例外是数字贷款机构快速增长的小额贷款业务，它们大规模应用自动化决策系统。

本节得出两点启示。第一，模型性能的宣称与观察到的结果之间存在差距，人们通过正当化、归责和自由裁量判断来应对这些差距，而自由裁量判断有时会凌驾于模型输出之上。第二，当政治或文化要求与数字借贷的市场逻辑发生冲突时，

[5]基尼系数也用于经济学，用于衡量某个人群内部收入或财富不平等的程度，取值范围从 0（完全平等）到 1（最大程度的不平等）。[6]基尼系数源自受试者工作特征曲线下面积（AUC）：基尼系数 = 2 × AUC - 1；基尼系数为 0 表示模型没有判别能力（相当于随机判断），而基尼系数为 1（或 100%）则表示完全区分。

违约率便会飙升，这引发了人们的疑问：信用评分能否在不系统性排斥借款人或依赖强制执行机制的情况下实现规模化运作。

6 讨论

上述记录的场景与故事，与此前人机交互和以人为中心的机器学习研究形成对话；这些研究聚焦于数据科学的社会技术与制度性工作，正是这类工作让机器学习系统得以运作 [71, 79, 80, 92],预测系统的技术局限[67, 121],各种隐私与问责考量的范围，以及它们如何引发各类干预 [26, 27, 115], 以及当数据和人工智能驱动系统遭遇现实世界时，人们所采取的自由裁量实践 [91, 95, 96]。在本研究中，我们观察了内罗毕信贷评分领域中初创企业及其他行动者的技术、制度与政治实践；在这一领域，“风险”是一种组织性技术与原则。通过民族志访谈和观察，我们追踪了对非传统或替代形式数据的获取范围扩大，如何预先塑造信用评分中风险的构建、合法化与评估方式。在我们的阐述中，风险被界定得愈发细致；这种界定一方面将个人与其更广泛的社会关系剥离，另一方面又利用这些关系重新界定并评估其信用worthiness。在第一部分实证分析中，我们展示了收集和分类替代数据的过程如何借助合法且往往有争议的方法，生成新型数据；然而，这些数据仍承袭了其所嵌入情境中的依赖关系。在第二部分中，我们考察了风险如何通过各种技术与社会定义加以构建，以支持模型设计与发展。这些扩展后的风险定义掩盖了某些形式的不确定性，进而引发了对模型性能与效能的担忧。在第三部分中，我们着重指出了人为自由裁量出现的政治与文化情境，尤其是在信用评分模型遭遇复杂现实情境之时。总体而言，我们的研究表明，风险与不确定性紧密交织；要让信用评分发挥作用，就需要持续进行对齐，即在模型与其声称要测量的世界之间开展双向转译。下文将讨论替代数据的施为性命名并产生风险分类的差异化影响；说明风险与不确定性的交织如何限制并挑战预测；并探讨对齐这一重要工作、它如何运作，以及它服务于谁。最后，我们将提出若干关于信用评分等高风险领域的研究与政策问题。

6.1 替代数据的施为性

正如我们的研究发现所示，替代数据预先塑造了算法信用评分的输出。这不仅是因为替代数据为模型所依赖的风险类别提供定义和更新所需的输入，也是因为这些数据从根本上具有施为性：它们会产生其所命名的效果 [5]。在我们的田野调查中，多次听到的一项主要论点是，扩大对替代数据的获取范围，将使初创企业和其他金融科技从业者能够减少欺诈和风险、提升自动化决策水平、拓宽数字贷款机构的利润空间，同时向原本被排除在外的群体提供信贷。尽管这些论点产生了不同的结果（下一节将对此展开讨论），但它们通过各种设备、分析技术和新形式的专业知识，以及对借款人和消费者行为所作的调整，催生了新的可读性机制，并实践了它们原本旨在衡量的世界。正如 Bowker [17] 所强调的，世界会在一定程度上被数据重塑：凡是被构建为可计数的事物，无论这种构建存在多大偏差，都会创造一个新世界；在这个世界里，这些计数会被常态化，甚至可能成为规范（例如将行为归为“良好”或“不良”、“高风险”或“安全”）。因此，施为性关乎谁或什么能够看见，以及谁或什么被看见，也关乎由此产生的影响。“谁或什么能够看见”包括促成数据收集、分析和实际运用的技术设备与机构实践，也包括获得这种新型（洞察）视野的行动者（从业者或公司）[84, 104, 113]。“什么被看见”则是指能够被分类和归类的特定数据类型（及其所代表的生活世界），包括那些被认为与预测模型的相关性较高或较低的数据 [52, 80, 103]。

纵观我们的研究发现，我们表明，替代数据的生产是一项社会与技术成就，依赖于：i）通过合法和/或可疑手段获取的大量数据；以及 ii）建立在对当地信贷文化与行为的了解之上的自由裁量判断。由于 Safaricom 对丰富的实时数据基础设施拥有垄断控制权，外部访问受到显著限制。在这一背景下，初创企业提出技术和道德层面的主张，为其收集由 Safaricom（以及包括银行和商业商店在内的其他行动者）生成的、可能无限量的轨迹数据的做法辩护；人们相信，这些数据能够捕捉个人或组织的性格与生活细节这一重要信息的丰富且可挖掘的象征性替代物 [127]。本研究中的初创企业通过技术变通方法和逆向工程构建这种替代数据——例如，研究发现中所描述的“短信抓取”实践。这些变通做法得以实现，依靠的是软件开发工具包和应用程序接口 [99]，它们使初创企业不仅能够自动收集其自有应用程序（即它们开发的贷款应用程序）中的数据，还能从更广泛的银行、信用参考局和金融科技应用程序中收集数据；对于这些应用程序，它们拥有“第三方”访问权限 [82]。“公私”合作以及非正式的数据共享协议等来源，使其能够访问范围更广、类型更多样的数据，从农业到体育博彩不一而足。尽管这些数据收集过程已被规范化为标准做法，但其运作处于合法性的边缘或超出合法性边界，使人们难以界定何为合乎伦理或规范，以及何为包容性做法或掠夺性做法。尽管使用了支持数据收集的“适当授权”系统（我们的几位参与者很快便援引了遵守肯尼亚《数据保护法》的必要性），但隐私侵扰和监控的程度——包括通过后台流程进行的隐私侵扰和监控——意味着，数据的收集和使用远远超出了其最初设定的目的（当然也超出了任何对情境完整性标准的合理适用范围 [63, 87, 118]）。

当分类和分析模式依据静态与动态等不同数据形式之间的区分时，它们拓宽了误解以及在信用评分中自动化或再生产偏见的可能性。当静态数据（规定个人的人口统计信息，如年龄、性别等）与动态数据混合时

，后者捕捉大量用户数据，包括实时客户交易、用户设备标识符，以及用户与其更广泛社交网络之间的联系 [5, 41]，，根据在美国语境中被称为“受保护属性”的信息进行推断和歧视的能力 [12] 会急剧扩大（在肯尼亚语境中，与这些属性相关的数据可通过基于应用程序的了解你的客户流程，甚至通过政府收费的综合人口登记服务（IPRS）API轻易获取 [19]）。由于替代数据促进了对用户日常行为和交易的动态及时间性层面的获取，我们的对话者将其视为正面（例如，活跃消费）和负面（例如，赌博）风险信号的来源。然而，所研究的初创企业在寻找风险信号方面如此执着，以至于良性且普通的行为和社会实践也被分类为“有风险”。例如，Moonshot 的数据科学负责人表示，他们为借款人的手机分配了唯一设备标识符，以分析是否有其他人使用同一部手机借款。通过将手机使用个体化，Moonshot 的风险评分将用户与具有社会和文化重要性的手机共享实践区分开来；这种实践在农村地区以及城市中的低收入人群中十分常见，而金融科技从业者声称要将这些人纳入金融体系。这些实践已在 HCI 和信息通信技术促进发展文献中得到广泛记录 [15, 21, 32, 123, 124]。当这些操作与生活世界中具有意义的结果联系起来时，正如从业者对“游戏”的担忧所表明的那样，用户开始重塑自身的行动和实践。渐渐地、微妙地，被计数的内容变成了具有重要意义的内容。

6.2 风险与预测的（社会技术）限度

我们的研究发现，风险不仅复杂且动态，还与不确定性密不可分（正如我们将在接下来的段落中展示的那样），这使得“风险”能否在实质上划定界限并转化为预测系统中的可计算对象变得更加复杂。数据科学团队中的对话者承认，尽管他们不断寻求越来越多的替代数据，机器学习仍然使用狭窄且具有历史局限性的（风险）数据快照进行预测。这是机器学习系统的根本设计（及其局限），多项关于情境化数据科学实践的研究对此进行了详尽记录，并展示了数据科学从业者如何认识和管理这些局限 [70, 79, 92]。通过持续改善数据质量和覆盖范围，并进行模型再训练，机器学习系统可以得到更新，达到“足够好”的程度 [14]。在我们的各个研究地点，替代数据被用作风险的代理指标——依据借款人的个人生活与社会生活——尤其是在缺乏可靠借款人信用数据的情况下（即所谓“信贷记录薄弱”或“信用隐形”的借款人）。此外，替代数据的动态性和时间性使数据科学团队能够更新其风险假设，并据此更新信用评分模型。正如我们的研究发现所示，用户所处情境中与心理与社会学层面相关的特征，是从静态与动态数据集中提取出来的。例如，研究人员设计特征来捕捉客户是否有储蓄习惯，或是否沉迷于体育博彩。虽然可以生成数百（甚至数千）个特征，但最终目标是将特征数量减少到少数几个，即 10 到 20 个。诸如风险降低之类的技术，例如相关性分析被广泛用于质疑可能产生虚假结果的人类直觉，并解释特定特征的重要性。尽管一些数据科学团队认为，储蓄等做法与信用worthiness呈正相关，而体育博彩与之呈负相关，但有几位参与者回忆起这些假设并不成立的情形（他们在这些“不可扩展的人工检查”中得出的许多见解，最终都成为专有秘密技术；他们表示，这些技术让他们比竞争对手更具优势）。

尽管数据科学实践中采用了各种变通做法（以及日常惯例）（相关性分析、可解释性工具、数据更新、模型再训练等），机器学习系统中风险与预测之间的关系往往并不稳固，尤其是在高风险决策领域。近年来，人机交互及相关领域的研究发现，预测系统往往会放大它们本欲缓解的那些风险 [106, 121]。其中一些研究质疑风险的认识论基础，强调时间性、系统性、环境性、社区层面等因素可能影响并扭曲风险预测结果 [106]。我们的研究结果表明，风险既掩盖了至少三种形式的不确定性，又与之纠缠在一起。

首先，认知不确定性与可知、因而可测之物相关 [59]。在许多方面，对替代数据的追求不仅是为了应对这种不确定性，也是为了拓展可知之物的边界。然而，尽管替代数据看似无所不包，基础设施依赖方面的缺口——其中一些已为人所知——仍阻碍了这些努力[50]。对于Safaricom等大型参与者而言，数据存储库极为碎片化，以至于他们几乎不可能在预测模型中利用这些数据：一位访谈对象透露，要切实完成此类数据的转化与使用，成本高得令人望而却步。与此相关，过度依赖SIM卡等身份基础设施来获取和核验借款人身份，使多家数字贷款机构暴露于这些数据基础设施的局限之中。虽然SIM卡使获取渠道更加普及，但其可互换性意味着数据档案经常被混淆，从而削弱最终的信用评分：当一张此前由贷款违约者持有的SIM卡在市场上流转时，新的SIM卡持有者可能会意外承担该贷款的负担（包括在信用参考局被列入黑名单的情形）。

其次，建模 不确定性涉及对机器学习模型设计的假设，以及对这些模型在现实世界中如何运作的主张 [61]。这种不确定性最明显地体现在一些方面：我们的对话者根据自己在数据中观察到的情况以及本地市场的运作方式作出建模决策。一个例子是使用信用额度而非信用评分（如美国及世界大部分地区的做法）。他们没有拒绝可疑的申请者，而是从一开始就为借款人设定专属信用额度（最低信用额度低至0.8美元），之后仅根据更多借款人行为数据进行调整。然而，将风险降至可控水平的过程，面临着数据集和评分模型中的动态变化与时间变化，当然也受到借款人日常行为的影响。

第三，情境 不确定性涉及对模型预测行使的自由裁量权 [95]。面对不断上升的不良贷款率和校准不佳的基尼系数，与

业务分析师和产品团队密切合作的对话者表示，他们会行使自由裁量权，以证明用其他类型的业务规则或政治理由补充预测评分的合理性。在信贷决策中使用这些系统的从业者会依据自身经验和专业知识作出判断，尤其是在信用模型未能达到既定承诺时。尽管模型开发者与信贷从业者之间的反馈循环探讨了改进的可能性（更多数据、混合评分等），但面对数字借贷领域相互重叠的工作流程以及不同的风险偏好，人们仍对能否彻底解决不确定性感到沮丧。

总的来说，数据科学和机器学习从业者认识到机器学习的设计局限，并运用各种技术和工具加以缓解。然而，风险与不确定性之间的关系，使人们难以实质性地声称模型具有预测准确率、有效性和一致性——尤其是在不完整数据、各异的建模假设以及从业者难以协调的自由裁量干预并存时。

6.3 对齐的对象、主体和时机是什么？

我们以 Fujimura [43] 和人机交互领域学者 [37, 39] 的研究为基础，将对齐界定为一种技术性与话语性工作，它通过共同面对的问题和利益，将实践的多个层面联结起来，促进这些层面的协调与协作。人机交互领域以往的研究主要考察了这种对齐的一个方向：通过可解释性、正当化和协商，让模型为组织与社会世界所接受 [11, 40, 92]。人工智能安全话语 [2, 51]（声称要将人类价值观和护栏纳入大语言模型——实际上就是让**“模型适应世界”**）沿用了这种单向框架，但将其视为一个规范性约束问题，而非对协调过程的分析性描述。与此相反，本文所说的对齐强调数据／模型与世界之间的双向转译，其中至少包含某种维度——正如我们已经看到的——即让**“世界适应模型”**。我们的研究发现，信用评分及其背后的技术、制度和世界观，被塑造成适应世界（安全、包容、负责任等）的形态；与此同时，世界本身也被塑造成适应信用评分的形态：借款人被规训为“良好”债务人，而手机共享等日常做法则被病理化为高风险行为。

这项工作部分得益于研究中各类行动者的努力、实践和认知，也得益于围绕通过信用评分技术实现信贷获取民主化的重要性的种种叙述。MoonshotExtenscore®的推出被誉为包容与责任的胜利，是数据科学家和资深分析师组成的技术团队的一项社会行动；他们通过创建一种混合评分，在风险管理方面胜过市场竞争者。这据称展现了替代数据的力量，能够为社会中服务不足和处于边缘地位的人群提供信贷。隐私法规（《数据保护法》及其配套法规）的制定，以及任命专员监督数字借贷中的数据违规行为（即数据保护专员办公室），被视为有助于提升算法信用评分的效能与安全性，并使其在市场上得到有序接纳。私营和公共部门开展的参与式流程，从用户调查、焦点小组，到纳入产品研发中的多元声音，再到中央银行邀请公众就提案草案和待出台法规发表意见，旨在为数字借贷产品和实践带来一定程度的监管与保障。

与此同时，我们看到世界正按照模型的“形象”被重塑的证据——行为、评估指标、实践与机会的调整，均源于信用评分日益扩大的存在及其重要性。由于不良贷款高企，借款人在许多情况下因被视为惯常违约者而受到谴责和指责（至少根据Hustler Fund场景中的PS所言）。金融科技公司和银行寻求新的机制，既纳入也排除潜在借款人，从而有效地为其服务筛选纳入和筛选排除新（及老）客户。这项工作在很大程度上具有规训性质，通过财务后果的威胁来强制得体行为与合规。借款人被要求成为“好”债务人：“你的行为就是你的财产。”虽然借款人特定数据本应与征信机构共享，以便在银行等受监管金融机构之间协调信用信息共享，但征信机构却常常被用作列入黑名单的场所，通过剥夺未来获得金融资源的机会来惩罚所谓的未偿还贷款情形（包括金额低于2美元的违规），进而对从住房到教育再到医疗保健等部门产生关键影响。这项工作的另一个方面还包括在不良贷款上升的情况下重新平衡贷款组合的表现，以及将失败合理化。例如，数字贷款机构依靠“好”客户来抵消“坏”借款人的影响（这些借款人反正会被收取高利率），从而维持其贷款账簿的增长。与此相关的是，尽管Hustler Fund录得市场最高违约率68.3%，其推动者仍为其不佳表现辩护，认为它是一项政策干预，而非以利润为导向的举措。

7 结论

在本文中，我们考察了在数据访问受限、信贷文化根深蒂固，以及肯尼亚政策与监管环境不断变化的背景下，机器学习和数据科学在开发信用评分系统过程中所开展的社会技术与制度性工作。我们指出，尽管初创企业、银行、电信公司、征信机构及其他金融科技从业者通过扩大替代数据的访问范围来重新界定风险，但风险从根本上与认识论、建模及情境形式的不确定性交织在一起，使模型在现实场景中的构建、部署和评估变得更加复杂。我们拓展了人机交互领域有关对齐的研究（并对人工智能安全论述提出了不同观点），指出信用评分是通过双向转译实现的：一方面让模型适用于现实世界，另一方面又重塑现实世界，使其符合模型假设。这种对齐工作带来了暂时的稳定和某种临时秩序，但对隐私、监控以及金融普惠条件造成的影响并不均衡。在这种双向转译的交汇处，存在着需要规范、实践乃至有时需要技术回应的矛盾。法律在保护基本隐私权的同时，信用信息框架却对替代数据的合法性（及其滥用）保持沉默，那么谁应负责弥补这些法律漏洞？围绕隐私的规范性问题应在何时、以何种方式与促进数据访问和普惠的发展主义愿景相平衡？在

缺乏惩罚性措施（例如在征信机构留下负面记录）的情况下，这些模型能否在复杂的现实情境中有效地大规模运行？或许最重要的是：什么构成风险，什么属于不确定性，又该由谁承担各自的代价与重担？

致谢

我们感谢 Miriam Fahimi 对早期稿件提出的深思熟虑的意见，感谢 Amiel Bize 和 Kevin Donovan 在整个项目期间持续投入，也感谢各位审稿人耐心而慷慨的反馈。我们还要感谢那些与我们分享时间、经历和专业知识的对话者。

参考文献

[1] Nikita Aggarwal。2023年。何时所有数据都是信用数据：消费信贷市场、技术发展与分配正义。博士论文。牛津大学，英国牛津。https://ora.ox.ac.uk/objects/uuid:975192f9-c333-4db7- 800f-6fcc0c417c7e[2] Shazeda Ahmed、Klaudia Jaźwińska、Archana Ahlawat、Amy Winecoff 和 Mona Wang。2024年。人工智能安全的领域构建与认识论文化。First Monday（4月 2024 10 5210 29 4 13626）。https://doi.org/ . /fm.v i 。[3] Rob Aitken。2017。。'所有数据都是信用数据'：构成无银行账户者。Competition & Change 21，4（2017年8月），274–300。https://doi.org/10.1177/ 1024529417712830[4] Saleema Amershi、Maya Cakmak、W. Bradley Knox 和 Todd Kulesza。2014年。将权力交给人民：人类在交互式机器学习中的作用。AI Magazine 35, 4 2014 105 120 10 1609 35 4 2513（），–。https://doi.org/ 。/aimag.v i 。电子版：https://onlinelibrary.wiley.com/doi/pdf/ 。/aimag.v i 。。[5] Louise Amoore。2013年。可能性政治：超越概率的风险与安全。杜克大学出版社，达勒姆。[6] Sareeta Amrute、Ranjit Singh 和 Rigoberto Lara Guzmán。2022年。关于全球多数世界中的人工智能：一处实证场域与一种立场的入门介绍。https: //doi.org/ 10 2139 4199467 . /ssrn。Julia Angwin、Jeff Larson、Surya Mattu 和 Lauren Kirchner。2016年。Machine Bias。https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing[8] Elizabeth A Ankrah、Kagonya Awori、Stephanie Nyairo、Mercy Muchai、Millicent Ochieng、Mark Kariuki、Gillian R Hayes 和 Jacki O'Neill。2025年。社会性是本性：社会架构如何塑造中小企业的工作，以及重新构想协作式人类—人工智能系统的思考。载于《2025年CHI计算机系统中的人为因素会议论文集》。ACM，横滨日本，1 19 10 1145 3706598 3715019 –。https://doi.org/ 。/ 。Arjun Appadurai。。[9] 金融机器中的幽灵。Public Culture 23，3 2011 517 539 10 1215 08992363 2011 1330399（），– 。https://doi.org/ 。/ -[10] 2016 Payal Arora。。数据金字塔底层：大数据与全球南方。International Journal of Communication 10，0（2016年3月），1681–1699。https://ijoc.org/index.php/ijoc/article/view/4297 编号：0。[11] Vijay Arya、Rachel K. E. Bellamy、Pin-Yu Chen、Amit Dhurandhar、Michael Hind、Samuel C. Hoffman、Stephanie Houde、Q. Vera Liao、Ronny Luss、Alek- sandra Mojsilović、Sami Mourad、Pablo Pedemonte、Ramya Raghavendra、John Richards、Prasanna Sattigeri、Karthikeyan Shanmugam、Moninder Singh、Kush R. Varshney、Dennis Wei 和 Yunfeng Zhang。2019年。一种解释并不适用于所有情况：人工智能可解释性技术的工具包与分类体系。https: //doi.org/10.48550/arXiv.1909.03012 arXiv:1909.03012 [cs]。[12] Solon Barocas 和 Andrew D. Selbst。2016年。大数据的差别影响。《加利福尼亚法律评论》104，3（2016年），671–732。[13] Rachel K. E. Bellamy、Kuntal Dey、Michael Hind、Samuel C. Hoffman、Stephanie Houde、Kalapriya Kannan、Pranay Lohia、Jacquelyn Martino、Sameep Mehta、Aleksandra Mojsilovic、Seema Nagar、Karthikeyan Natesan Ramamurthy、John Richards、Diptikalyan Saha、Prasanna Sattigeri、Moninder Singh、Kush R. Varsh- ney 和 Yunfeng Zhang。2018年。AI Fairness 360：用于检测、理解和缓解不良算法偏见的可扩展工具包。https: //doi.org/10.48550/arXiv.1810.01943 arXiv:1810.01943 [cs]。[14] Paula Bialski。2024年。Middle Tech：软件工作与"足够好"的文化。第 36 10 2307 9021578 卷。普林斯顿大学出版社。[15] 2016 https://doi.org/ 。/jj。Nicola J. Bidwell。。去殖民化人机交互与交互设计话语：规划 AfriCHI 时的一些考量。XRDS: Crossroads，《ACM学生杂志》 22, 4 2016 22 27 10 1145 2930884（6月），– 。[16] https://doi.org/ 。/ Daniel Björkegren、Joshua E. Blumenstock 和 Samsun Knight。2020年。防操纵机器学习。https://doi.org/10.48550/arXiv.2004. 03865 arXiv:2004.03865 [econ]。[17] Geoffrey C. Bowker。2000年。生物多样性，数据多样性。《科学社会研究》30，5（2000年10月），643–683。https://doi.org/10.1177/030631200030005001 [18] Geoffery C. Bowker 和 Susan Leigh Star。2000年。《梳理万物：分类及其后果》。麻省理工学院出版社，美国马萨诸塞州剑桥。[19] Keith Breckenridge。2019年。《"关于肯尼亚人的唯一真实信息来源"的失败：NDRS、抵押品之谜与 Safaricom 垄断》。非洲研究 78, 1 2019 91 111 10 1080 00020184 2018 1540515（1月），–。10 1080 00020184 2018 1540515 https://doi.org/ 。/。。电子版：https://doi.org/ 。/。。[20] Benjamin Amani Brühwiler。2015年。《欠债与放贷的道德观：肯尼亚达累斯萨拉姆 Kariakoo 的信贷、债务与城市生活》。博士论文。密歇根州立大学。[21] Jenna Burrell。2010年。《评估共享接入：社会平等与乌干达农村地区移动电话的流通》。《计算机媒介传播期刊》15卷2期（2010年），230–250。https://doi.org/10.1111/j.1083-6101.2010.01518.x[22] Bruce G. Carruthers。2013年。《从不确定性走向风险：信用评级案例》。《社会经济评论》11卷3期（2013年7月），525–551。https://doi.org/10. 1093/ser/mws027[23] Stevie Chancellor、Eric P. S. Baumer 和 Munmun De Choudhury。2019年。《以人为中心的机器学习中的"人"是谁：从社交媒体预测心理健康的案例》。《ACM人机交互汇刊》 3, 2019 1 32 10 1145 3359249 计算机支持的协同工作（11月），–。https://doi.org/ 。/ Jean Comaroff 和 John L. Comaroff。《南方理论：或欧美如何演变并走向非洲》。Paradigm出版社，博尔德（科罗拉多州）、伦敦。[25] Tamara Cook 和 Claudia McKay。2015年。《M-Shwari 如何运作：迄今为止的故事》。技术报告10。CGAP 和肯尼亚金融部门深化组织，华盛顿哥伦比亚特区。https://www.cgap.org/sites/default/files/Forum-How-M-Shwari-Works-Apr-2015.pdf[26] A. Feder Cooper、Emanuel Moss、Benjamin Laufer 和 Helen Nissenbaum。2022年。《算法社会中的问责：机器学习中的关系性、责任与稳健性》。载于2022年ACM公平、问责与透明度会议论文集。ACM，韩国首尔，864–876。https: //doi.org/ 10 1145 3531146 3533150 . / . Amanda Coston、Anna Kawakami、Haiyi Zhu、Ken Holstein 和 Hoda Heidari。2023年。《从有效性视角评估数据驱动决策算法的合理使用》。IEEE计算机学会，690–704。https://doi. org/ 10 1109 54575 2023 00050 . /SaTML . . Christopher Csíkszentmihályi、Jude Mukundane、Gemma F. Rodrigues、Daniel Mwesigwa 和 Michelle Kasprzak。2018年。《可能性的空间：撒哈拉以南非洲的技术创新政治经济》。载于《2018年CHI计算机系统中的人为因素会议论文集》（CHI '18）。计算机协会，美国纽约州纽约，1–13。https://doi. org/ 10 1145 3173574 3173880 . / . Jay L. Cunningham、Sydney T. Nguyen、Julie A. Kientz 和 Daniela Rosner。2022年。《文化的代价：Cash App 与美国黑人社区金融普惠情况分析》。载于交互系统设计会议。ACM，加拿大线上活动，612 628 10 1145 3532106 3533569 –。[30] 2025 https://doi.org/ 。/。 Ankolika De、Shaheen Kanthawala 和 Jessica Maddox。《谁能发声？揭穿关于招募和让非WEIRD人群参与研究的"难以触及"迷思》。载于《2025年ACM公平、问责与透明度会议论文集》（FAccT '25）。计算机协会，美国纽约州纽约，855–867。https: //doi.org/ 10 1145 3715275 3732055 . / . Alain Desrosières。[31] 1998 《大数字的政治：统计推理史》。哈佛大学出版社。[32] Jonathan Donner。2015年。《接入之后：包容、发展与更具移动性的互联网》。麻省理工学院出版社。OCLC：907512564。[33] Kevin P. Donovan。2024年。《金钱、价值与国家：东非的主权与公民身份》。剑桥大学出版社，剑桥。https://doi.org/10.1017/ 9781009501385[34] Kevin P. Donovan 和 Aaron K. Martin。2014年。《非洲SIM卡注册的兴起：监管变革的新兴动态》。First Monday（2014年1月）。https://doi.org/ 10 5210 19 2 4351 . /fm.v i . Kevin P. Donovan 和 Emma Park。2019年。[35] 《硅谷草原的永久债务》。《波士顿评论》（2019年9月）。https://www.bostonreview.net/articles/perpetual-debt-silicon-savannah/[36] Paul Dourish。2007年。《像界面一样观察》。载于《第19届澳大拉西亚人机交互会议：趣味用户界面》（OZCHI '07）论文集。计算机协会，美国纽约州纽约，1–8。https://doi.org/ 10 1145 1324892 1324894 . / . Paul Dourish。《设计的魅力与匮乏：设计文化与文化中的设计》。2021 。《人机交互》36卷1期（2021年1月），52–72。https://doi.org/10.1080/07370024.2018.1469410 _电子版：https://doi.org/10.1080/07370024.2018.1469410。[38] Miriam Fahimi 和 Katharina Kinder-Kurlanda。2023年。《价值物质性中的摩擦：关联透明度、算法与信用评分》。《数字文化与社会》 9, 2 2023 141 160 10 14361 2023 0208（12月），–。[39] https://doi.org/ 。/dcs- - Miriam Fahimi、Mayra Russo、Kristen M. Scott、Maria-Esther Vidal、Bettina Berendt 和 Katharina Kinder-Kurlanda。2024年。《公平性中的协调工作与反复调试》。《ACM人机交互汇刊》8卷，计算机支持的协同工作第2期（2024年11月），434:1–434:23。https://doi.org/10.1145/3686973

[40] Fair Isaac Corporation。2018年。《机器学习与 FICO®评分：有助于贷款机构和消费者的机器学习创新演进》。白皮书。11页。https://www.fico.com/en/resource-access/download/6359 [41] 2024 Marion Fourcade 和 Kieran Healy。《序数社会》（第1版）。哈佛大学出版社，剑桥。[42] Nicolas Friederici、Michel Wahome 和 Mark Graham。2020年。《非洲的数字创业：一个大陆如何摆脱硅谷的深远影响》。麻省理工学院出版社，剑桥。[43] Joan H. Fujimura。1987年。《在癌症研究中构建"可行"问题：阐明对齐》。《科学社会研究》17卷2期（1987年），257–293。https: //www.jstor.org/stable/284931 [44] 2012 David Graeber。《债务：最初的5000年》（十周年纪念版）。Melville House，纽约州布鲁克林。[45] Prince K. Guma。2020年。《转型中的城市基础设施不完整性：内罗毕移动时代的情景》。《科学社会研究》50卷5期（10月 2020 728 750 10 1177 0306312720927088），–。[46] 1990 Ian Hacking。《驯服偶然》。剑桥大学出版社，剑桥 [英格兰]; 纽约。[47] Yousif Hassan。2025年。《作为国家建设实验的机器学习：非洲的人工智能与发展》。载于《牛津机器学习社会学手册》，Christian Borch 和 Juan Pablo Pardo-Guerra 编。牛津大学出版社，https://doi.org/ 。/ 。/oxfordhb/。。[48] 2016 Mikella Hurley 和 Julius Adebayo。《大数据时代的信用评分》。《大数据》18（2016年）。[49] Azra Ismail、Divy Thakkar、Neha Madhiwalla 和 Neha Kumar。2023年。《人工智能公共卫生呼吁：民族志视角》。《ACM人机交互汇刊》7卷，计算机支持的协同工作第2期（2023年9月），1–26。https: //doi.org/ 401 420 10 1093 9780197653609 013 27 –。[50] Edwards、Geoffrey C. Bowker 和 Cory P. Knobel。2007年。《理解基础设施：历史、启发式方法与网络基础设施政策》。First Monday 12卷6期（2007年6月）。https://doi.org/10.5210/fm.v12i6.1904[51] Jiaming Ji、Tianyi Qiu、Boyuan Chen、Jiayi Zhou、Borong Zhang、Donghai Hong、Hantao Lou、Kaile Wang、Yawen Duan、Zhonghao He、Lukas Vierling、Zhaowei Zhang、Fanzhi Zeng、Juntao Dai、Xuehai Pan、Hua Xu、Aidan O'Gara、Kwan Ng、Brian Tse、Jie Fu、Stephen Mcaleer、Yanfeng Wang、Mingchuan Yang、Yunhuai Liu、Yizhou Wang、Song-Chun Zhu、Yike Guo、Yaodong Yang 和 Wen Gao。2025年。《人工智能对齐：当代综述》。《ACM计算调查》58卷5期（11月 2025 132 1 132 38 10 1145 3770749），：–：。https://doi.org/ 。/ [52] 2019 Anne Jonas 和 Jenna Burrell。《摩擦、蛇油与怪异国家：网络安全系统可能通过区域封锁加剧全球不平等》。《大数据与社会》6卷1期（2019年1月），2053951719835238。https://doi.org/10.1177/ 2053951719835238[53] Maximilian Kasy 和 Rediet Abebe。2021年。《算法决策中的公平性、平等与权力》。载于《2021年ACM公平、问责与透明度会议论文集》。ACM，加拿大线上活动，576–586。https://doi.org/10.1145/3442188.3445919[54] Anna Kawakami、Amanda Coston、Hoda Heidari、Kenneth Holstein 和 Haiyi Zhu。2024年。《自上而下研究公共部门人工智能：权力关系网络如何塑造围绕人工智能设计与使用的机构决策》。http://arxiv.org/abs/2405. 12458 arXiv:2405.12458 [cs]。[55] 肯尼亚人权委员会。2025年。《辜负奋斗者：Hustler Fund 是披着经济赋权外衣的重大政治噱头》。技术报告。肯尼亚人权委员会，肯尼亚内罗毕。37页。https: //khrc.or.ke/wp-content/uploads/2025/08/FAILING-THE-HUSTLERS-3-5.pdf[56] Morris Kiruga。2020年。《这款借贷应用在你逾期还款前对你关怀备至。之后，羞辱便开始了》。https://restofworld.org/2020/okash-microlending-public-shaming/ [57] Barbara Kiviat。2019年。《美国的信用评分》。《经济社会学：欧洲电子通讯》21卷1期（2019年），33–42。[58] Leora Klapper、Dorothe Singer、Laura Starita 和 Alexandra Norris。2025年。《2025年全球金融包容性数据库：数字经济中的连接与金融普惠》。华盛顿特区：世界银行。https://doi.org/10.1596/978-1-4648- 2204-9[59] Frank H. Knight。1921年。《风险、不确定性与利润》。Hart, Schaffner & Marx，波士顿，马萨诸塞州。https://www.loc.gov/item/21012860/ 最初作为 Hart, Schaffner & Marx经济学奖论文集的一部分出版。[60] Lindah Kotut。2025年。《条款与条件（不）适用：理解移动金融服务设计中的剥削差异》。载于 AfriCHI '25：非洲人机交互会议。https://doi.org/10.48550/ arXiv.2507.10970[61] Tianqi Kou、Dana Calacci 和 Cindy Lin。2025年。《问责死区：为何机器学习研究中的社会主张应予以阐明并加以辩护》。https://doi.org/10.48550/arXiv.2508.08739 arXiv:2508.08739 [cs]。[62] Paul Langely。2014年。《为创业者配备资源：消费信贷与信用评分》。《消费、市场与文化》17卷5期（2014年9月），448–467。https: //doi.org/ 10 1080 10253866 2013 849592 . / . . Josh Lauer。《信用可靠：美国消费者监控与金融身份史》。哥伦比亚大学出版社。https://doi.org/10.7312/ [63] 2017 laue16808 [64] Yeuk Yu Lee 和 Rosanna Bellini。2025年。《权衡之中：协作金融技术的洞见》。载于《2025年CHI计算机系统中的人为因素会议论文集》（CHI '25）。计算机协会，美国纽约州纽约，1 19 10 1145 3706598 3713119 –。https://doi.org/ . / . Q. Vera Liao 和 Kush R. Varshney。。以人为中心的可解释人工智能（可解释人工智能）：从算法到用户体验。https://doi.org/10.48550/arXiv. 2110.10790 arXiv:2110.10790 [cs]。[66] Chuncheng Liu。2019年。《中国的多个社会信用体系》。https://doi.org/ 10.2139/ssrn.3423057[67] Lydia T. Liu、Inioluwa Deborah Raji、Angela Zhou、Luke Guerdan、Jessica Hull-man、Daniel Malinsky、Bryan Wilder、Simone Zhang、Hammaad Adam、Amanda Coston、Ben Laufer、Ezinne Nwankwo、Michael Zanger-Tishler、Eli Ben-Michael、Solon Barocas、Avi Feller、Marissa Gerchick、Talia Gillis、Shion Guha、Daniel Ho、Lily Hu、Kosuke Imai、Sayash Kapoor、Joshua Loftus、Razieh Nabi、Arvind Narayanan、Ben Recht、Juan Carlos Perdomo、Matthew Salganik、Mark Sendak、Alexander Tolbert、Berk Ustun、Suresh Venkatasubramanian、Angelina Wang 和 Ashia Wilson。2025年。《弥合社会系统中的预测与干预问题》。https://doi.org/ 10 48550 2507 05216 2507 05216 . /arXiv. . arXiv: . [ ] cs . [68] Kristian Lum 和 William Isaac。2016年。预测并服务？《Significance》13卷5期（2016年），14–19。https://doi.org/10.1111/j.1740-9713.2016.00960.x _电子版: https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1740-9713.2016.00960.x。[69] Eric Gwandega Magale。2024年。《迈向金融普惠？肯尼亚合作型与数字借贷模式研究》。博士论文。比勒陀利亚大学，南非比勒陀利亚。[70] Momin M. Malik。2020年。《机器学习的局限层级》。https: //doi.org/10.48550/arXiv.2002.05193 arXiv:2002.05193 [cs]。[71] Yaoli Mao、Dakuo Wang、Michael Muller、Kush R. Varshney、Ioana Baldini、Casey Dugan 和 Aleksandra Mojsilović。2019年。《数据科学家如何与领域专家开展科学合作：找到正确答案，还是提出正确问题？》《ACM人机交互汇刊》3卷GROUP期（12月 2019 237 1 237 23 10 1145 3361118），：–：。https://doi.org/ . / [72] 1995 George E. Marcus。。《世界体系中的民族志：多点民族志的兴起》。《人类学年度评论》24卷，1995年第24卷（10月 1995 95 117 10 1146 24 100195 000523），–。[73] 2020 https://doi.org/ . /annurev.an. . . Vidushi Marda 和 Shivangi Narayan。。《新德里预测性警务系统中的数据》。载于《2020年公平、问责与透明度会议论文集》（FAT* '20）。计算机协会，美国纽约州纽约，317 324 10 1145 3351095 3372865 –。https://doi.org/ . / . Rafe [74] 2024 Mazer 和 Seth Garz。。《快速增长与缓慢政策：肯尼亚数字信贷十年》。《牛津经济政策评论》40卷1期（2024年3月），82–103。https://doi.org/10.1093/oxrep/grad055[75] Ninareh Mehrabi、Fred Morstatter、Nripsuta Saxena、Kristina Lerman 和 Aram Galstyan。2022年。《机器学习中的偏见与公平性综述》。《计算机调查》54卷6期（2022年7月），1–35。https://doi.org/10.1145/3457607[76] Jacob Metcalf、Ranjit Singh、Emanuel Moss、Emnet Tafesse 和 Elizabeth Anne Watkins。2023年。《将算法带上法庭：一种关系性算法问责方法》。载于《2023年ACM公平、问责与透明度会议》。ACM，美国伊利诺伊州芝加哥，1450–1462。https://doi.org/10.1145/ 3593013.3594092[77] Kepha Muiruri。2025年。《数字贷款机构面临压力，违约率最高达40%》。https://www.businessdailyafrica.com/bd/economy/heat-on-digital-lenders-as-default-rate-hits-up-to- 40 4978996 [78] 2010 pc- Michael Muller 和 S. Kogan。。《人机交互与计算机支持的协同工作中的扎根理论方法》。（2010年1月）。[79] Michael Muller、Ingrid Lange、Dakuo Wang、David Piorkowski、Jason Tsay、Q. Vera Liao、Casey Dugan 和 Thomas Erickson。2019年。《数据科学工作者如何与数据打交道：发现、采集、整理、设计与创建》。载于《2019年CHI计算机系统中的人为因素会议论文集》（CHI '19）。计算机协会，美国纽约州纽约，1–15。https://doi.org/ 10 1145 3290605 3300356 . / . Michael Muller 和 Angelika Strohmayer[80]。2022年。《数据科学中的遗忘实践》。载于《2022年CHI计算机系统中的人为因素会议论文集》（CHI '22）。计算机协会，美国纽约州纽约，1 19 10 1145 3491102 3517644 –。https://doi.org/ . / . Collins [81] W. Munyendo、Yasemin Acar 和 Adam J. Aviv。2022年。"非常时期需要非常手段"：肯尼亚用户对手机贷款应用的担忧。载于《2022年IEEE安全与隐私研讨会》（SP）。IEEE，美国加利福尼亚州旧金山，2304 2319 10 1109 46214 2022 9833779 –。https://doi.org/ . /SP . . Grace Mutung'u 和 Kevin Muchwat。。《肯尼亚数字借贷应用的隐私与数据保护实践》。《知识产权与信息技术法杂志》（JIPIT）1卷（2021年6月），131–169。https: //doi.org/ 10 52907 1 1 68 . /jipit.v i . Daniel [82] 2021 Mwesigwa 和 Christopher Csikszentmihályi。2024年。《空中／时间旅行：重新思考全球人机交互中的挪用与电子交换的未来》。载于《CHI计算机系统中的人为因素会议论文集》。ACM，美国夏威夷州檀香山，1–21。https://doi.org/10.1145/3613904.3642590 [83]

[84] Naja Holten Møller、Claus Bossen、Kathleen H. Pine、Trine Rask Nielsen 和 Gina Neff。2020年。谁在做数据工作? 《interactions》27卷3期（2020年4月），52–55。https://doi.org/10.1145/3386389[85] Simone Natale、Federico Biggio、Payal Arora、John Downey、Riccardo Fassone、Rafael Grohmann、Andrea Guzman、Emily Keightley、Deqiang Ji、Vincent Obia、Aleksandra Przegalinska、Usha Raman、Paola Ricaurte 和 Eduardo Villanueva-Mansilla。2025年。《全球人工智能文化》。《Communications of the ACM》（2025年8月）。https://doi. org/10.1145/3722547 在线优先发表。[86] Gina Neff。2012年。《风险劳动：创新产业中的工作与风险负担》。麻省理工学院出版社。https://doi.org/10.7551/mitpress/9780262017480.001. 0001[87] Helen Nissenbaum。2019年。《数据食物链上下游的情境完整性》。《法律理论探究》20卷1期（2019年3月），221–256。https://doi. org/ 10 1515 2019 0008 . /til- - Rebecca Njuguna 和 Karen Sowon。2021年。海报：[88]替代信用评分文献范围综述。载于ACM计算与可持续社会特别兴趣小组会议。ACM，线上活动（澳大利亚），437–444。https: //doi.org/ 10 1145 3460112 3471972 . / . Rachel O'Dwyer。。算法在评估信用评分时，重蹈[89] 2018了人类一个世纪前的覆辙。https://qz.com/1276781/algorithms-are-making-the-same-mistakes-assessing-credit-scores-that-humans-did-a-century-ago 栏目：市场。[90] Emma Park 和 Kevin P. Donovan。2023年。《隐私、匮乏与人：数据、债务与基础设施化人格》。载于《非洲的技术转译》，Richard Rottenburg、Faeeza Ballim 和 Bronwyn Kotzen（编）。第1卷：度量指标。博睿出版社。https://doi.org/10.1163/9789004678354_004 栏目：非洲的技术转译。第1卷：度量指标。[91] Samir Passi 和 Steven Jackson。2017年。《数据视野：学会透过算法抽象观察》。载于《2017年ACM计算机支持的协同工作与社会计算会议论文集》（CSCW '17）。计算机协会，美国纽约州纽约，2436–2447。https://doi.org/10. 1145/2998181.2998331[92] Samir Passi 和 Steven J. Jackson。2018年。《数据科学中的信任：企业数据科学项目中的协作、转译与问责》。《ACM人机交互汇刊》2卷CSCW期（2018年11月），1–28。https: // doi.org/ 10 1145 3274405 . / Martha Poon。。《消费金融中的系统性风险》。[93] 2010《Limn》1期（2010年12月），38–38。https://limn.press/article/systemic-risk-in-consumer-finance/ [94] Theodore M. Porter。2020年。《统计思维的兴起，1820–1900》（新版）。普林斯顿大学出版社，新泽西州普林斯顿。OCLC：on1197737788。[95] Juho Pääkkönen、Matti Nelimarkka、Jesse Haapoja 和 Airi Lampinen。2020年。《以官僚制为视角分析和设计算法系统》。载于《2020年CHI计算机系统中的人为因素会议论文集》。ACM，美国夏威夷州檀香山，1 14 10 1145 3313831 3376780 –。https://doi.org/ . /[96] . Rida Qadri 和 Catherine D'Ignazio。2022。。像司机一样观察：劳动者如何修复、抵抗并强化平台的算法愿景。《大数据与社会》9, 2 2022 20539517221133780 10 1177 20539517221133780 （7月），。[97] 2021https://doi.org/ . / Farah Qureshi、Stephen C. Rea 和 Kristin N. Johnson。。《（解）构金融普惠主张：人工智能在美国和肯尼亚消费信贷市场中的应用——评估人工智能的承诺与风险》。《国际法与比较法杂志》8, 2 2021 405 434 = 8 （），–。https://heinonline.org/HOL/P?h hein.journals/jintcl &i =409[98] Divya Ramesh、Vaishnav Kameswaran、Ding Wang 和 Nithya Sambasivan。2022年。《平台与用户的权力关系如何塑造算法问责：即时贷款平台与印度财务困窘用户案例研究》。载于《2022年ACM公平、问责与透明度会议》。ACM，韩国首尔，1917–1928。https://doi.org/10.1145/3531146.3533237[99] Joel Reardon、Álvaro Feal、Primal Wijesekera、Amit Elazari Bar On、Narseo Vallina-Rodriguez 和 Serge Egelman。2019年。《泄露数据的50种方式：探究应用程序如何绕过Android权限系统》。载于第28届USENIX安全研讨会。美国加利福尼亚州圣克拉拉，603–620。[100] Richard Rothstein。2018年。《法律的颜色：政府如何隔离美国的被遗忘历史史》（2018年Liveright平装本首版）。Liveright出版公司（W.W. Norton & Company旗下），纽约、伦敦。[101] Ananya Roy。2010年。《贫困资本：小额信贷与发展之道》（第一版）。劳特利奇出版社，纽约。https://doi.org/ 10 4324 9780203854716 . / Minna [102] 2017Ruckenstein 和 Natasha Dow Schüll。。《健康数据化》。《人类学年度评论》46卷，2017年第46卷（2017年10月），261–278。https: //doi.org/ 10 1146 102116 041244 . /annurev-anthro- - Nithya [103] 2021Sambasivan。。像全球南方的数据集那样观察。《interactions》28卷4期（2021年7月），76–78。https://doi.org/10.1145/3466160[104] Nithya Sambasivan、Shivani Kapania、Hannah Highfill、Diana Akrong、Praveen Paritosh 和 Lora M Aroyo。2021年。"人人都想做模型工作，而不是数据工作"：高风险人工智能中的数据级联。载于《2021年CHI计算机系统中的人为因素会议论文集》。ACM，日本横滨，1–15。https://doi.org/10.1145/3411764.3445518

[105] Devansh Saxena、Karla Badillo-Urquiola、Pamela J. Wisniewski 和 Shion Guha。2020年。《美国儿童福利系统中所使用算法的以人为本审查》。载于《2020年CHI计算机系统中的人为因素会议论文集》（CHI '20）。计算机协会，美国纽约州纽约，1 15 10 1145 3313831 3376229 – 。https://doi.org/ 。/ 。[106] Devansh Saxena、Erina Seh-Young Moon、Aryan Chaurasia、Yixin Guan 和 Shion Guha。2023年。通过对儿童福利个案记录进行计算叙事分析，重新思考算法系统中的"风险"。载于《2023年CHI计算机系统中的人为因素会议论文集》（CHI '23）。计算机协会，美国纽约州纽约，1–19。https://doi.org/ 10 1145 3544548 3581308 。/ 。[107] 1996AnnaLee Saxenian。《地区优势：硅谷与128号公路的文化与竞争》。哈佛大学出版社。Google图书ID：gnh2Rb1rcMIC。[108] Andrew D. Selbst、Danah Boyd、Sorelle A. Friedler、Suresh Venkatasubramanian 和 Janet Vertesi。2019年。社会技术系统中的公平性与抽象。载于《公平、问责与透明大会论文集》。ACM，美国乔治亚州亚特兰大，59 68 10 1145 3287560 3287598 – 。[109] 2021https://doi.org/ 。/ 。Phoebe Sengers、Kaiton Williams 和 Vera Khovanskaya。《推测与发展设计》。《ACM人机交互汇刊》5, 1 2021 121 1 121 27 10 1145 3449195 CSCW（4月），：–：。[110] 2019https://doi.org/ 。/ Alain Shema。《使用有限手机数据进行有效信用评分》。载于《第十届信息通信技术促进发展国际会议论文集》（ICTD '19）。计算机协会，美国纽约州纽约，1 11 10 1145 3287098 3287116 – 。[111] 2024https://doi.org/ 。/ 。Benjamin Shestakofsky。《创业公司幕后：风险投资如何塑造工作、创新与不平等》。加州大学出版社，美国加利福尼亚州奥克兰。[112] Parker Shipton。2010年。《跨文化的信贷：非洲的农民、金融家与误解》。耶鲁大学出版社，美国康涅狄格州纽黑文。https: //doi.org/ 10 12987 9780300162929 。/ Ranjit Singh 和 Steven Jackson。2021年。[113]以基础设施的视角看待：低分辨率公民与Aadhaar身份识别项目。《ACM人机交互汇刊》5，CSCW2（2021年10月），315:1–315:26。https: //doi.org/10.1145/3476056[114] Jessie J Smith、Anas Buhayh、Anushka Kathait、Pradeep Ragothaman、Nicholas Mattei、Robin Burke 和 Amy Voida。2023年。公平性的多重面貌：探索多利益相关方小额贷款推荐的制度逻辑。（2023年）。[115] Daniel Solove 和 Danielle Keats Citron。2021年。隐私伤害。《乔治·华盛顿大学法学院教师出版物及其他作品》（2021年1月）。https://scholarship.law.gwu.edu/faculty_publications/ 1334[116] Skyler Speakman、Srihari Sridharan 和 Isaac Markus。2018年。用于手机信用评分的三种人群协变量偏移。载于《第一届ACM SIGCAS计算与可持续社会会议论文集》。ACM，美国加利福尼亚州门洛帕克和圣何塞，1 7 10 1145 3209811 3209856 – 。[117] 2025https://doi.org/ 。/ 。Tatiana Thieme。《拼搏式城市主义：在内罗毕谋生》。《全球化与社群》第36卷。明尼苏达大学出版社，美国明尼苏达州明尼阿波利斯。Edward Kahuthia Murimi作后记。[118] Nanna Bonde Thylstrup、Kristian Bondo Hansen、Mikkel Flyverbom 和 Louise Amoore。2022年。机器学习系统中的数据再利用政治：理论化再利用纠缠。《大数据与社会》9，2（2022年7月），20539517221139785。https://doi.org/ 10 1177 20539517221139785 。/ Radha Upadhyaya、Keren Weitzberg 和[119] Linda Bonyo。2025 年。数字信贷提供商、监管框架与结构性权力：肯尼亚数字小额信贷监管的个案研究。《金融与社会》（2025年4月），1–23。https: //doi.org/ 10 1017 2025 4 。/fas。/ Tessie Waithira。《无魂的债务引擎》。《[120] 2024Kernel》杂志4（2024年）。https: //www.kernelmag.io/4/debt-engines [121] Angelina Wang、Sayash Kapoor、Solon Barocas 和 Arvind Narayanan。2024年。反对预测优化：论优化预测准确率的决策算法的正当性。《ACM负责任计算期刊》1, 1 2024 1 45 10 1145 3636509 （3月），–。[122] 2009https://doi.org/ 。/ 世界银行。《营商环境报告：在困难时期推进改革》。世界银行。https://doi.org/ 10 1596 978 0 8213 7961 5 。/ - - - - Susan P. Wyche 和 [123] 2012Laura L. Murphy。《"中国制造的死机"手机脱离电网：调查并设计面向非洲农村地区的手机使用方案》。载于《交互系统设计会议论文集》（DIS '12）。计算机协会，美国纽约州纽约，186–195。https://doi.org/10. 1145/2317956.2317985[124] Susan P. Wyche、Nightingale Simiyu 和 Martha E. Othieno。2019年。理解肯尼亚农村妇女的手机使用：一种可供性方法。《移动媒体与传播》7，1（2019年1月），94–110。https: // doi.org/ 10 1177 2050157918776684 。/ Malte Ziewitz。重新思考游戏：网络搜索引擎中优化的伦理实践。[125] 2019《科学社会研究》49，5（2019年10月），707–731。https://doi.org/10.1177/0306312719865607

[126] Julie Zollmann。2020年。《微薄度日：应对现代肯尼亚的经济匮乏》。Practical Action Publishing，沃里克郡拉格比。 https://doi.org/10.3362/9781788531207

[127] Shoshana Zuboff。1988年。《智能机器时代：工作的未来与权力》。Basic Books, Inc.，纽约，美国。