

数字是骨头：为什么拆掉风险评分的所有理由，却仍然拆不掉它

评分不靠测得更准存活，而靠成为决断者在制度压力下得以站立的支撑性器官

高鹏 著 · SDE 学员 · 算法治理与制度分析 · 2026年7月26日

摘要 关于累犯风险评分的批评已积累数十年：它存在统计偏差、引发自证预言、固化社会不平等、侵蚀罪刑法定原则。这些批评在智识上几乎已被赢下——然而，评分系统不仅存活，还在扩张。本文诊断这一“拆不掉”的困境，指认既有批评共享一个未经审查的先验：它们都假设评分首先是一个“测量”工具，只是测偏了或被社会误用了。本文撤销这一假设，从“决断者与判断重负之间的张力”这一更底层的矛盾出发，重新结算评分系统的实质。评分之所以拆不掉，不是因为它测得更准，而是因为它从诞生之初，其核心功能就已从“提供信息”悄然滑向“提供支点”——一种让决断者得以在制度压力下站立的支撑性器官。本文提出“痕迹存在论依赖”来描述这根骨头的运作机制：它不靠“测准”存活，而靠把整个社会的日常决断寄生在自己的输出上，让个体的数字痕迹替代其完整存在成为制度操作的本体，并以此锁死自身被拔除的可能性。本文进而以美国联邦量刑指南的历史变迁为一个完整案例，追踪骨头如何生成、如何被依赖、以及拔骨之痛的实质所在；在此过程中分析骨头排斥与吸纳两种命运路径的生成逻辑；论证这一概念与布迪厄的误识、福柯的规训、埃斯波西托的免疫、以及“逃避自由责任”这一古老命题之间的分离线；并与当代算法治理文献对话，明确本文的独特贡献在于将分析焦点从“算法如何塑造被治理者”转向“算法如何支撑并暴露治理主体的结构性软弱”。最后，本文给出证伪条件：若将评分用于资源支持而非自由剥夺、且切断其与量刑的直接因果链后，既有的惩罚系统在更长时间尺度内失去公众对其必要性的共识，则本文对“骨头”的诊断成立。

关键词：累犯风险评分；痕迹存在论依赖；算法治理；决断卸责；量刑指南

一、问题叠着问题：为什么拆掉评分反而更难了？

关于累犯风险评分的批评，过去三十年间已经积累成一个令人敬畏的武库。它被指出内嵌种族与阶层偏见：一个在贫穷街区长大、年少时被捕过的人，终其一生背负着那些静态变量锁定的“高风险”分数，而模型从未问他长大的街区有多少家银行、多少可及的合法工作。它被揭示在运行自证预言：高分延长刑期，长期监禁切断社会联系，出狱后的合法路径随之收窄——于是设计者指着再犯数据说，看，我们预测得多准。它被指控侵蚀了现代刑法的一个奠基性原则：我们惩罚的是已然的行为，不是未然的风险——当评分把一个人的家庭背景、药物史、性格问卷得分换算成更长的刑期时，它在法理上已越过了那条线。

这些批评在智识上是强有力的。它们不缺乏论证的力量，不缺乏数据，不缺乏法理，甚至不缺乏体制内部分改革者的善意。但它们缺乏一样东西：对“为什么拆不掉”这个问题的诚实面对。

在北美，COMPAS及其同类量表在被反复曝光、起诉、学术围剿后，仍然每年被用于数十万人的量刑、假释与监管级别评估。它们有时被调整——加入一两个新变量，换一套训练数据，改进“公平性”指标——但他们的核心地位不仅未被撼动，反而从刑事司法扩展到了儿童保护、精神卫生强制收治、校园威胁评估等新场域。我们面对的不是一个原地固守的堡垒，而是一个边挨批评边繁衍的有机体。迄今为止，批评者的典型策略是不断升级批评的精度：找出更隐蔽的偏差，揭示更复杂的反馈回路，呼唤更根本的系统性替代方案。这些努力并非徒劳，它们持续暴露着评分系统的认知缺陷与伦理债务。但没有一种理论出来回应一个更令人坐立不安的可能：也许这套东西的韧性，根本不在它“测得多准”里面；也许它之所以经得起所有批评，不是因为批评的力度还不够，而是因为批评打在了它的铠甲上——而那副铠甲本身，并不是它最重要的器官。

本文试图提供一个不同的诊断。我们不再追问“评分是准还是偏”，不再追问“它是否制造了自证预言”，甚至不再追问“它是否违背了罪刑法定原则”。我们将从一套更底层的张力出发——不是测量者与被测量者之间的张力，而是社会里每个决断者与他必须背负的决断责任之间的张力——来重新看清楚：累犯风险评分在当代社会里到底在干什么？它不是一把尺。它是一根骨头。一根长在社会身体里的、一旦拔掉整个社会都会发现自己不知道该怎么站的骨头。[1]

二、撤掉老地基：评分不是一把尺

在进入诊断之前，有必要把本文与既有批评在存在论上的分岔点交代清楚——不是靠援引一套外来语汇，而是就着这个领域自身不曾被质疑的操作来说。

所有针对累犯风险评分的既有批评，无论其立场多么激进，共享着同一个“老地基”：它们认为评分首先是一个测量工具，只是测错了，或者被用错了，或者被嵌入了不公平的社会结构里。这是批评者们站立的操作台：我们指责一把尺刻度歪了——这是统计偏差批评；我们指出这把尺在测量时把被测物压变形了——这是自证预言批评；我们抗议这把尺被用来打人而不是量东西——这是罪刑法定批评。所

有这些批评动作，都是在和这把尺较劲——较它的准头，较它被使用的方式，较它被分配到的权力。"评分是一把尺"这个预设，从未被这些批评本身撤销过。

然而，如果评分并不是一把尺呢？如果"作为测量工具"只是它的功能外衣，而它实际上在做的事，是另一件批评从未瞄准的事呢？

让我们从一个最朴素的现象入手。任何一个被COMPAS或同类量表评定为"高风险"的个体，他接下来遭遇的所有社会决断——法官的量刑裁量、假释委员会的释放评估、观护官的监督级别——都能一致地、流畅地引用这个分数来为自己的决定背书。法官不再需要说出"我内心觉得他不可信"，他说"客观风险评估显示高风险"。假释委员会不再需要说出"我们担心放他出去有事"，他们说"精算评估不支持释放"。这整个链条的流畅性，是"尺子"这个词解释不了的。一把尺如果刻度歪了，用尺的人会争吵——谁用错了，谁看偏了，谁应该校正。但评分系统引发的不是争吵，是默会。它激活的不是对"准不准"的争论欲，而是一个决断者把自己决断的"声音"切给它、然后自己退到后面的那一步轻省。

要理解这一步轻省是怎么变成整个系统的默认操作的，我们必须看到：测量与支点，从一开始就是两种不同的功能。早期的累犯风险预测工具——比如二十世纪二三十年代的假释决策表——在设计者的蓝图中确实首先是"尺"：它试图用精算方法替代当时被认为过于随意、充满偏见的临床判断，为假释官提供更可靠的预测信息。在这个阶段，工具的使用者仍在做判断；工具只是提供参考数据。然而，随着二十世纪下半叶刑事司法系统面临的双重压力——一方面是社会对犯罪控制的焦虑不断升高，另一方面是司法系统内部的效率与标准化压力持续加剧——这套工具的角色发生了静默的漂移。法官和假释官需要的，逐渐不再是一份可供参考的"数据"，而是一个可以在判决书、假释报告中直接引用的"客观依据"。这个漂移在制度层面的完成，是1980年代量刑指南改革的一个意外后果：当量刑被纳入一张可计算、可比较、可上诉审查的网格时，风险评分就从一个"辅助信息"变成了一种"合规资源"——它帮助决断者在制度的格子里找到那个不会被上级推翻的位置。

累犯风险评分的真正产品，从来不是"对累犯风险的准确预测"。它的真正产品，是让整个社会所有需要做出艰难决断的人，都可以把自己的决断挂在一根"数字骨头"上——然后用自己的嘴说出骨头的话，用自己的手下写骨头的签名。法官的量刑、假释官的评估、雇主的背景审查——这些动作本来每一次都需要执行者自己去面对一个具体的人，自己去找支点，自己去承担判断的沉重。骨头替他们省掉了这一步。它给每个决断者提供了一个现成的、可被引用、可被追溯、可被免责的支点。决断者不再需要用自己的去担保那个判断；骨头的客观性替他担了。这根骨头不靠"测得准"来存活。它靠"所有人都在用"来存活。用的人越多，骨头就越硬；骨头越硬，它就越是不能被某一份批评报告拔掉——因为你拔的不是一个错误的工具，你拔的是整个社会已经长在上面的一整套日常决断的支点。那把尺歪了，换一把就是。骨头歪了，身体得先学会不用骨头站——这比换尺难一万倍。

我们将这种机制命名为：痕迹存在论依赖。[2]"痕迹"指的是：评分系统产生的不是一个活着的人的故事、理由与面容，而是一串连续的、可在数据库之间流转的数字痕迹——分数、等级、标签。"存在论"在此的用法是：这些痕迹一旦生成，就被相关决断者当作那个人的"所是"本身来使用——不是当作"他的部分信息"，而是当作"他这个人"可以被决断系统直接处理的那个本体形式。所谓"依赖"，指的是社会决断链的各个环节在日常操作中已经长在这些痕迹上，离开了它们，决断者会发现自己不会做判断——不是技术上不会，而是伦理上不敢、组织上不被允许、时间上不被给予。

三、骨头如何自己喂饱自己：一场三方互锁

如果说上一节给出了"骨头"的命名并追溯了它从"尺"到"骨"的历史漂移，本节要做的是让这个命名动起来——描述那颗"数字骨头"的运转生态。我们关心的是：它是如何让所有想拔掉它的人都使不上力的？答案不在于评分系统本身的强悍，而在于它把三种不同的人锁进了彼此喂料的互锁闭环：决断者——他们需要甩掉决断的重负；被评分者——他们被迫反向依赖分数的痕迹来换取一丝喘息；社会旁观者——他们需要被数字安抚，以确认系统仍在运转、"坏人"仍在被管住。三方各自从这根骨头里提取的东西不一样，但三方都是骨头的喂料者。

第一角：决断者的轻省寄生。想象一个法官的量刑决策过程。他面前有两套逻辑。一套逻辑是面对面的、道德判断的：他需要理解这个被告的生平脉络，掂量他的悔意是真是假，感受他在法庭上的态度是挑衅还是恐惧，去判断判决对他未来的影响——并且，最后在判决书上签下自己名字的时候，他只是这个判断的唯一责任承担者。这套逻辑很重。另一套逻辑是精算的、数字的、匿名的：被告的COMPAS暴力累犯风险被评为"高"，该辖区的量刑指南允许将此作为加重因素。法官在判决理由中写入"基于客观风险评估"，这等于把判决的伦理重量从自己肩上卸下一部分，放在了一台机器上。他不再是一个孤零零站着的人，他身后站着一个算法，而算法不说话，算法只显示结果。

这恰恰是批评最难撼动的东西。当一篇学术论文告诉这位法官"COMPAS在种族之间有预测偏差"时，法官听到了——但他听到的，和他每天在法庭上感受到的，是两个层面的东西。偏差是统计学内部的纠葛；而轻省，是一个人每天要连续做出十几个会实质改变他人命运的决定的唯一可持续方式。我们需要诚实面对的是：批评在告诉法官"你的尺子不准"的那一刻，已经假设了法官想要一把更准的尺子。但法官要的不是更准的尺子。法官要的是骨头。准不准，从来不是骨头的第一功能。给你一个支点，让你能站住，才是骨头的第一功能。哈考特在其对精算方法的系统批判中，曾准确地指出精算逻辑并不仅仅"预测"犯罪——它通过改变执法资源的分配与监控的强度，内生地制造着它声称只在"预测"的犯罪分布。[3]本文进一步指出：正是因为这一内生机制使得预测的"准确性"在统计上永远可以被自我印证，它才

为决断者提供了一种特别牢固的支点——不是因为测准了，而是因为无论结果如何，都可以在事后被模型自身的逻辑解释为"合理的"。

第二角：被评分者的反向寄生。 高分者并不是只承受分数——他们也学会了使用分数。在假释听证会上，一个被评定为"高风险"的囚犯，他无法让自己的申诉被听到，除非他能提供另一套同样客观、同样可被数字表述的证据：他已经完成了多少小时的心理辅导，他在愤怒管理课程里的出勤率，他的狱内违规记录是否为零。他想松动这根骨头对他的压迫的唯一方式，不是砸碎骨头——他砸不碎。他只能用骨头自己的逻辑来为自己争取一点微弱的位移：不是"这个评分错了"，而是"请看，我在这套评分的另一页上，有一些可以加分的痕迹"。这使得被评分者反过来成为骨头的维护者。他无法承受拆掉骨头的后果——因为骨头一旦消失，他连唯一的申辩语言都会失去。他会突然掉进一个他完全不熟悉的重力场：无法再用"我已修满课程"说话，而必须以一个人的完整存在去说服另一个人的心。

第三角：社会旁观者的安全感提取。 那些从不直接接触刑事司法系统、却支撑着这个系统之政治正当性的广大公众，他们是如何感知"恶人"是不是被管住了的？他们看不见每个被告的生平，听不懂法庭上的交叉辩论，不可能跟踪每一个出狱者的生活轨迹。他们能看见的，是数字：累犯率是升了还是降了；这个罪犯的风险评分是多少；那个监狱的再犯率是百分之几。数字供他们摄取安全感。如果明天所有累犯风险评分被废除，法官恢复纯粹的临床综合判断，公众不会觉得"判断终于恢复了人味"——公众会觉得"判断又退回了不可靠的、因人而异的拍脑袋"。他们会要求恢复数字，因为数字给他们的不是准确，是安心的形式：有一个表，有一个分数，它看起来控制得住。这根骨头因此不仅被决断者所依赖，它还被社会的合法性感知所依赖。

这三方——决断者要甩掉的重量、被评分者要抓住的申辩语言、社会旁观者要提取的安全感——不是三个各自独立的现象。它们就是骨头的血液循环系统。批评者的每一篇文章，每一次敲打，都打在了"准不准"上。但骨头的本质就不在准不准。它只需要让所有参与这场社会仪式的玩家，都能找到自己在这个仪式里的位置。

四、骨头雕出的两类命运：两种被逼出的路

数字骨头不仅支撑着决断者，它还在持续雕琢被它钉住的人。我们通常认为，评分制造出的是"被排斥者"。事实比这更分化。骨头制造的不是两类人，而是两种命运的产出路径。我们关心的是：一道初始分数，如何在个体与制度的反复碰撞中被逐步硬化，最终把同一个人推向截然不同的生存轨迹。

先看排斥的路径。一个人最初被标记为高风险，这个痕迹进入制度后，并不直接"导致"他被排斥。它与制度的碰撞是一步一步发生的。第一次碰撞：带着高风险分数的囚犯参加假释听证，假释委员会引用评分拒绝释放。"高风险"痕迹在这一次碰撞中被激活并强化——这不是因为评分在一个孤立事件中"算准了"，而是因为拒绝假释本身延长了他的监禁时间、切断了他建立合法社会联系的机会，从而实质性地提高了他未来再犯的概率。第二次碰撞：几年后他再次出现在假释委员会面前，此时他的"高风险"痕迹已经不仅是当初那个分数——他在狱中的每一次违规记录、每一次拒绝参加课程的选择，都被制度以骨头的语言重新归档为"持续危险"的证据，进一步压低了释放可能。第三次、第四次碰撞接连发生。每一次，他都面对同一个困境：他如果不会用骨头的语言说话，他无法让制度听到任何有利于他的理由；但如果他试图用自己的语言说话——"我感觉自己已经变了"、"我想回家照顾母亲"——制度没有可以处理这些陈述的接口。他被堵住的不是是一条路，而是他被制度"听到"的可能性本身。最终，当他累积了足够多次的拒绝、足够厚的负面痕迹，他就被推到了防火墙那边——不是被某个人的恶意驱逐，而是被一次次制度碰撞里无法翻译成痕迹语言的生命叙事所逐出。他的命运不是被"诊断"出来的，而是被一步一步逼窄的：每一步碰撞都缩小了他下一次选择的半径，每一步制度回应都加固了痕迹的预言力，直到他没有路可以再走。

再看吸纳的路径。另一些处境下的被评分者——他们或许分数略低，或许更早学会了使用制度内的语言——在碰撞中走上了一条表面上更好、实质上同样由骨头铺设的路。最初，他发现，要换得假释委员会的一次点头，他需要修够心理辅导的课时；要换得观护官的一笔正面报告，他需要完成职业培训的出勤率。他开始积累这些可被量化的痕迹。第二次碰撞时，他带着这些痕迹去谈判：我用这一叠证书换取一个宽松一点的监管级别。制度接受了。第三次碰撞，他再次带着新的痕迹去交换更多一点的自由度。代价是什么？代价是，他逐渐不再问"我想成为什么样的人"。他只在制度能读懂的格子里行动，只积累制度能归档的痕迹。他不再用自己的声音说话——他学会了用骨头的语法说话。他不是被排斥的，他是被吸纳的；而这种吸纳的代价，是一整个主体性的格式化。海绵吸满了体制给的水，但海绵永远只能是体制给它的那种形状。

是什么让同一个体在这两条路径之间被推向这一条而非另一条？决定性因素并非个体的"性格"或"改造意愿"。推手首先在于他最初被骨头锁定的位置——高风险与中低风险初始痕迹，从一开始就设定了制度愿意为这个个体开放多少可积累正向痕迹的接口。其次在于他所拥有的、可被制度识别的"反向操作"资源——一个在狱中有稳定家庭支持、能够提供工作录用意向书的囚犯，比一个举目无亲的人更容易在制度的接口上找到可以"填入"正向痕迹的格子。第三，在于他在与骨头的每一次互动中，是学会了使用它的语言，还是试图用自己的语言对抗它——而每一次对抗的失败，都在教给他同一个教训：不用骨头说话，你就没有说话的位置。

五、拔骨之痛：不是改参数，而是失去"用分数说话的权利"

至此，我们已能回答那个起始的悖论：为什么批评赢了道理却输掉了现实？因为批评一直在用测量学的语言打一场存在论战争。它以

为自己在攻击一个认识论对象——一个预测模型，准确率堪忧，偏差严重。但它其实面对着一个存在论器官——一根长在社会身体里的骨头，不靠准确率而活，靠三条血管的持续供养维持存活。

要拆掉一根骨头，有两类手术。一类是修复术：打磨骨刺，矫正骨形，增加几个变量使其“更公平”。这是所有既有改革——算法公平、透明审计、加入动态变量——的实质。它们想做的换一把尺。它们能缓解一些局部症状，但它们永远触不到骨头的本体，因为修复术在动刀的那一刻就默认了骨头必须留在体内。修复术越精细，骨头越被认为“可以修好”，它就越不可能被拿掉。

另一类是移除术：不再假设它需要留在体内，问的问题是——如果这根骨头被拿掉，身体能不能学会用肌肉自己站？

要回答这个问题，我们不能只靠抽象推演。美国联邦量刑指南的局部松动，提供了一个未经规划却极具诊断价值的准自然实验。1984年的量刑改革法案建立了强制性的联邦量刑指南，将量刑从法官的自由裁量转变为一个由“犯罪等级”和“犯罪历史分数”严格计算出的格子系统。这不是简单地“使用”一个评分工具，而是将整个量刑决策变成了一个由数字痕迹驱动的机械化程序。2005年，联邦最高法院在布克案中裁定，强制性量刑指南违宪——法官必须“参考”指南，但不再受其强制约束。

布克案之后发生了什么，是这个故事最关键的一章。它实际上进行了一次局部的、非自愿的“拔骨手术”。法官们被还回了裁量权，指南从“必须服从的规则”变成了“可供参考的信息”。然而，实证研究追踪了此后十余年的量刑实践，发现了一个极具张力的现象。一方面，判决确实出现了更大的差异性：一些法官在特定类型的案件中更倾向于判低于指南建议的刑期。但另一方面，指南的“参考”地位并未让法官们整体性地摆脱对数字框架的依赖。许多法官继续将指南的计算结果作为量刑的锚点，在此基础上上下浮动；判决书中的量刑理由仍然大量引用指南的计算步骤，只是换了一种更柔性的措辞。联邦量刑委员会自身的年度报告持续追踪布克案后各巡回法院的“指南遵从率”——这个指标本身的存在就说明了问题：即便在强制性被废除之后，整个司法系统仍然需要一种方式去衡量法官“有没有好好用尺子”。[4]

这个案例暴露了拔骨手术的真正难度。当一个已经长在社会身体里的支点被撤销强制力时，决断者并不会自动“学会自己站立”——他们会去寻找同一根骨头的柔性版本，继续靠着它。指南从规则变成参考，但它没有消失；它只是换了一种方式支撑着决断。而那些确实开始更独立地进行裁量的法官，他们付出了什么样的代价？他们需要在上诉审查中为自己的“偏离指南”提供更详尽的书面理由；他们的判决更容易被质疑为“主观”或“不一致”；他们在同僚中更显得“标新立异”。布克案释放出的，不仅仅是裁量自由；它同时释放出了一个制度性的冷风：当你不再靠骨头站的时候，你需要独自面对全部的风。

移除术的真正门槛，从来不在技术可行性。你可以把风险评分从量刑和假释决策中拿掉，只保留在心理健康支持、职业培训、住房帮扶等资源分配端——技术上，这是可以做的。门槛在别处。门槛在于：当我们把它从剥夺端拿掉的那一刻，所有曾经依赖它来为自己决断背书的法官和假释官，会瞬间失去一个他们已用了半辈子的语言。他们必须重新学会用“我判断”而不是“数据显示”来开口说话。那意味着他们必须重新承受道德判断的完整重量。更关键的是，他们需要对抗的不仅是自己的习惯，而是一整套让“自己判断”变得危险、让“引用骨头”变得安全的制度环境——案量压力、上诉风险、行政提速要求，已经把“做出一个无法被算法背书的裁量”异化为一种职业冒险。这才是拔骨最难下的那一刀：不是对算法的修复更困难，是让决断者在一整套不鼓励他们自己判断的制度里，重新承担判断本身——太沉重。

更诚实地说：即便社会真的开始了“学习用肌肉站立”，也没有一个“学会之后”的黄金状态在终点等着。用肌肉站立，不是一个一次性习得的技能，而是一个需要反复承受的、更容易疲劳、更容易摔倒、且永远不会像靠骨头那样省力的常态。拔骨不是通向一个更纯真的决断状态，而是通向一个需要不断在矛盾中重新配置决断责任的过程——每一次配置都可能失败，每一次失败都可能诱使骨头重新长出。

六、这根骨头在别处：校园里的那一道早期分数

如果我们的分析只停在刑事司法，就无异于在承认一个奇特的存在论特权：风险评分的骨化机制，只当人们被叫成“罪犯”时才生效。事实显然不是这样。痕迹存在论依赖是一种弥散性的社会器官形态，它在多个领域已长出类似的结构。信用评分让一个因医疗账单破产的人被标记为高风险借款人，这个痕迹让他借不到正常的钱，只能去边缘借贷，陷入更深的债务，每一次逾期都喂养着那个评分，让他的“财务人格”被评分允许他进入的金融生存空间所塑造。精神医学里的诊断标签——一个十五岁被诊为“品行障碍”——进入他的病历，成为他在每一个精神卫生机构之间的移交痕迹，在事后每一次医患接触中都先于患者本人而被读到，然后反过来决定了这次接触里患者的行为被怎样归档、怎样理解。

但这些领域的逻辑与刑事司法高度同构。本节挑选一个不同于刑事司法、却能最清晰地展示这根骨头如何“把人发生潜能的路切掉”的案例，做一个完整的微型解剖——不是描述现象，而是拆开那个生成链条：一道数字痕迹，是如何通过制度节点的逐级传递、通过日常微观互动的反复加固，最终把一个孩子的潜能从“尚未发生”变成“不再可能发生”的。

设想一个在小学三年级标准化阅读测试中得分偏低的孩子。这个分数本身只是一次性的事件，但它进入学校的教务系统后，开始了一道痕迹的骨化过程。第一步：分数触发了分班程序中的一条自动规则——他被分入“基础加强班”，这个班的课程进度比普通班慢两个教学单元，阅读材料的难度降了一级，课堂讨论的时间被压缩、换成了更多的语法练习。这道痕迹在此刻完成了它的第一次制度硬化：从一个测试分数，变成了一种被组织化了的、低期望的课程环境。

第二步才是更关键的——痕迹通过日常微观互动进入认知层，锁死教师动态更新判断的可能性。新老师在开学前已经拿到班级名单，上面有每个学生的测试纪录与标注。从这一刻开始，那道分数就在教师的认知中发挥着三重效应。第一是归类效应：教师需要在一个有三十个学生的班级中快速分配注意力资源，分数替她做了预先筛选——那些被标注为“需要加强”的孩子，在未经任何当面互动之前，就已经被放在了一个“可能需要更多辅导但进步空间有限”的认知格子里。第二是诠释锁定效应：当这个孩子在正式课堂上第一次答错问题时，教师的这个观察被前一道痕迹吸收、解读为“他果然是低水平”的确认；而如果另一个从未有过低分标注的孩子答错了同样的问题，教师更可能解读为“他今天状态不好”或“这道题太难”。同一个行为，被两条不同的痕迹路径导入两套不同的制度反应。第三是自我保护效应：教师也是在一个被评估的系统里工作的。年终评估时，她需要解释为什么班上的部分学生成绩低于年级平均——如果她曾把一个低分记录的孩子当作“高潜力”来要求，万一期末他成绩仍然偏低，她需要为这个“判断失误”承担责任。反过来，如果她从开学就把他定位为“需要加强”，她的记录和操作都是合规的、可被追溯的、可被解释的。这些微观操作汇聚在一起，构成了一个闭合回路：教师不是主动决定“放弃期待”，而是制度的日常运转——课程表的安排、督导检查的标准、年终评估的问责——已经把“对一个低分标注的孩子予以更高期待”变成了一个对她的职业而言成本更高、风险更大的选择。

第三步，孩子在课堂上从同伴的眼神、老师分发的练习卷、自己被提问的频率中逐渐感觉到：自己不属于“那种会读书的孩子”。他开始不再举手。他开始告诉自己，自己不擅长阅读。这不是一个心理学的“自我实现预言”在无中介地运作。它是一个制度性的生成过程：低期望的课程环境降低了他接触复杂文本的机会，教师诠释锁定效应减少了他获得启发式反馈的次数，他的阅读能力确实没有获得在另一个轨道里可能获得的成长——而这个结果，在两年后的下一次标准化测试里，会被当作“证据”，证明三年前的那道早期分流是准确的。那道分数从一开始就被设计成“测量”，但它实际上完成了从测量到制造的闭环。他不是“不适合念书”。他是一道早期痕迹通过制度内的归类、诠释锁定与自我保护三重微观操作，逐步被剥夺了让潜能发生出来的机会。

七、撞一撞敌意最近邻：这根骨头的诊断到底多了什么

任何一个足够有力的诊断，都必须能说明它和此前最靠近的那些诊断之间的差异——不是用“它们都不全，我比它们更全”这种空话，而是指认出一条清晰的分离线：那些已经占住位的概念，在哪一个点上无法覆盖本文要说的东西？本文的核心命名“痕迹存在论依赖”及其核心隐喻“数字骨头”，必须正面回应至少四个已经在这个议题附近深耕已久的思想资源：布迪厄的“误识”、福柯的“规训”、埃斯波西托的“免疫”、以及一个更古老的命题——“逃避自由责任”。此外，它还需要与当代算法治理研究做出区分，明确自身的理论定位。

与布迪厄“误识”的分离。布迪厄的“误识”描述的是：社会行动者将权力结构的任意性误认为理所当然的自然秩序，从而在无意识中再生产了不平等。这句话听起来似乎也能解释评分系统的韧性——法官误以为评分是“客观的”，公众误以为分数代表“真实的危险”。但“痕迹存在论依赖”抓住的不是“误”，是“靠”。然而，要真正完成这一分离，我们不能拿布迪厄理论最单薄的版本来当对手。布迪厄思想中最锋利的那层意思，不是“行动者太天真、被蒙蔽了”——而是行动者对整套游戏规则的大前提不加追问。在这个意义上，一个布迪厄主义者完全可以把本文所说的“寄生”收编为一种更深层误识的运作：法官之所以依赖骨头，是因为他误识了“判决必须由一个非人化的支点来支撑才具有合法性”这一整套司法场域的入场规则本身——他以为这是理所当然的，而不追问它是否必须如此。

这个更尖锐版本的“误识”必须被正面接住。本文的回应是：即便如此，“误识”仍然解释不了一种关键情形。设想一位法官，她在某个时刻识破了这套游戏规则——她意识到，判决的合法性并不真的需要爬在一个数字工具上，它也可以来自判决书里对个体处境更充分、更诚实的说理。她识破了误识。但她明天开庭时，面前仍然是排满的案卷；她写出一个不引用精算工具的判决，仍然需要在上诉审查时面对比标准化判决高得多的被推翻风险；她的结案率仍然是绩效考核的指标。在这种情况下，她是否会停止引用评分？误识理论没有资源处理这个问题。它是一个关于“认知”的理论，不是一个关于“制度压力下行动支点”的理论。本文与布迪厄的分离线因此应该这样表述：误识讲的是认知被游戏规则的黑会前提所封闭；痕迹存在论依赖讲的是，即使那个认知封闭被打开了，行动者仍然可能在制度压力下无法放弃那个支点。你可以停止误识，但你无法停止需要支点——不是因为你的认知没打开，而是因为你所处的制度环境没有给你不用支点站立的实际空间。

与“逃避自由责任”命题的分离。本文诊断所面临的最有力的潜在反驳，不是任何一个具体的学术概念，而是一个古老的、几乎是常识性的命题：人类从来都在逃避判断的重负——从神明裁判、陪审团裁决到精神病学鉴定，每一次不过是换一种卸责工具。骨头只是最新、最有效的一种。这个反驳如果成立，会直接削弱本文的原创性——你不是发现了一根新的“骨头”，你只是给人类一个永恒的老毛病起了一个时髦名字。

这个反驳必须被正面回应，而且回应它恰恰是让“痕迹存在论依赖”的不可还原性最清晰地现身的途径。之前的卸责工具——神谕、陪审团裁决、精神鉴定——共同特征在于，它们产出的是一段叙事、一个事件、或一份可以被另一份同行鉴定质疑的判断。神明裁判说出的是一个可以被祭司重述的神意；陪审团给出的是一个不可以被追问推理过程的裁决——但它是一个一次性的、不可被自动迁移至后续法律程序中的事件；精神鉴定师出具的报告，在法庭上可以被另一份鉴定报告质疑，在交叉询问中可以被拆解。所有这些前数字时代的卸责工具，产出的都是可以重新诠释的东西。数字骨头产出的是什么？是一段可跨数据库流转、可被各制度环节自动读取和引用、且在存在论上被操作它的系统当成个体“所是”本身的数字痕迹。它不是关于那个人的一段叙事，它在制度的日常操作中就是那个人——不是隐喻意义上的“就是”，而是操作意义上的：假释委员会在做出决定时能接触到的、能处理的、能在表格里打勾的，不是这个人的生故事，是他的

分数、等级、风险标签。这不是从“卸责工具”到“新卸责工具”的量变。这是从“卸责工具”到“卸责存在论”的质变——人的痕迹替代了人本身成为制度操作的主体，而这个替代一旦完成，就无法被任何面对面的叙事推翻，因为制度已经不设处理那些叙事的接口了。

与福柯“规训”的分离。福柯的规训权力通过持续的监视、检查、标准化训练，把个体塑造成“驯顺的身体”——主体被权力从外部制造出来。这个框架在描述被评分者（特别是“海绵”）时似乎相当好用：被评分者不断修课、攒分、更新档案，不就是在进行一场自我规训吗？但规训理论漏掉了一个关键的传动轴：决断者的角色。在福柯的规训全景里，监视者本人也处在权力网络之中，但监视者并不“依赖”那个被规训者来分担他自己的决断重压。本文描述的机制恰好相反：法官不是权力金字塔上一个无名的监视员，他是一个每天要亲手签署判决、为判决承担被上诉风险、被公众审视的具体的、焦虑的决断者。他引用评分，不是因为他在执行一套规训程序，而是因为他主动甩掉一块他自己扛不住的石头。规训讲的是“权力通过身体运作”；骨头讲的是“权力通过让决断者能用掉决断的身体来运作”。分离线：规训中被盯住的是被治理者；骨头中被盯住的是治理者自己的软弱。

与埃斯波西托“免疫”的分离。埃斯波西托从生命政治传统中提炼出“免疫”逻辑：社会为了保护自身整体安全，必须纳入一小部分被控制的“威胁物”（比如疫苗里的灭活病毒、刑罚系统里的罪犯），通过这种纳入来激活整个共同体的防御机制。风险评分在这个框架中可以被解读为一种免疫工具——它识别出高风险个体，通过监禁或监控将他们“纳入”可控范围，从而保护社会整体。“免疫”讲的是系统通过纳入来保护自己。但本文追踪的机制走了另一条路：系统通过输出一个可被各环节寄生的支点来维持运转。排斥与吸纳的实质，不是他们的“威胁”被纳入管理，而是他们的“痕迹”被各环节用来作为卸责的支点。免疫理论中，被纳入者是功能性的——他们被用来激活保护；骨头理论中，被钉住者是结构性的——他们被用来撑住决断。分离线：免疫吸纳威胁以保卫全体；骨头输出痕迹以绑定决断。

与算法治理研究的对话。当代科学技术研究与法律社会学中关于“算法治理”的丰富文献，同样触及了算法如何重塑制度逻辑、如何生产主体性。这些研究揭示了算法如何将个体转化为可计算、可排序、可干预的“数据化客体”；它们分析了“算法作为行动者”如何在制度场景中行使原本属于人类的裁量权；它们批判了“预测性正义”话语如何将刑法的时间箭头从过去转向未来，从而侵蚀了罪责原则。本文的诊断与这些研究共享同一个问题域，但在分析焦点上做出了一次明确的转移。既有研究的核心追问可以概括为：算法对被治理者做了什么？本文的追问则是：算法对治理者做了什么？本文的论点是：评分系统之所以不可替代，不是因为它在塑造被治理者方面特别高效（尽管它确实如此），而是因为它在支撑治理者方面特别“好用”——它让治理者不必独自背负判断的道德重量。在这一视角下，算法批判的通常结论——“算法把人物化”——被补充了新的维度：算法之所以能把人物化，首先是因为治理者在制度压力下需要一种可以把人物化的工具来让自己站稳。这不是在否定既有研究的价值，而是在它的基础上打开一个新的政治与伦理问题域：关于治理者自身的结构性软弱，关于决断责任在制度环节中的分配与逃避，关于“客观性”作为一种制度性卸责机制的存在论根源。

完成了这四场对质与一场对话之后，“痕迹存在论依赖”宣称的存在论收益可以这样收束：它抓住的，是评分系统最不可替代的那个功能——不是认识（为决断者提供正确信息），不是规训（塑造驯顺的身体），不是免疫（隔离威胁），也不是逃避自由责任的一个新案例。它是支撑——为所有需要做出艰难道德决断的人，提供一个可以让他们的决断站起来的、客观的、可追溯的支点；并且，这一支点的产出物不是可以被重述的叙事，而是可以被制度当作个体“所是”本身来操作的数字痕迹。布迪厄、福柯、埃斯波西托、“逃避自由责任”的古老命题、以及当代算法治理研究各自从不同的角度照亮了这个系统的某些侧面，但它们都没有瞄准“决断时刻的制度性卸责”与“制度环节对痕迹支点的结构性依赖”这一组关系。这组关系，正是“痕迹存在论依赖”要命名的那个剩余物。

八、重新追问“邪恶”：它不再是内心的事，而是骨头挤出的一个位置

让我们回到那个古老的词：邪恶。如果要为前面的分析补上最直接、也最容易被回避的一笔追问，就是这一笔：为什么在一个号称拥抱数据客观性的时代，社会反而比任何时候都更依赖制造这样一类被数字钉死的“不可接纳者”？他们不完全是经典犯罪学所描绘的那种邪恶——不是内心的黑暗，不是道德的堕落，不是反社会人格的冷酷算计。他们是被数字骨头从社会身体里挤出来的一个必然位置。

这个位置是怎么被发生出来的？让我们倒退几步。在现代刑事司法系统尚未被精算逻辑渗透的时期，社会如何感知“恶人已被管住”？在公开处决和肉体酷刑的时代，惩罚的可见性——刑场上的惨叫声、枷号示众的羞辱——为公众提供了直接的“安全感提取”：他们亲眼看见惩罚发生了。当刑罚退入监狱高墙之后，这种可见性消失了。社会需要新的方式来确信“惩罚系统仍在运作”——由此生出了犯罪统计、再犯率、以及后来的风险评分。这不是一个“社会有需求所以发明了工具”的单向因果。这是一个渐进的、在实践中被逼出来的生成过程：每一次公众对犯罪的道德恐慌、每一次媒体对司法“软弱”的指控、每一次立法机关对法官“自由裁量权过大”的抨击，都在制度中积累压力，要求惩罚系统拿出可以被数字衡量的“成效”证明。风险评分就是在这些压力的反复撞击下，从精算学的边缘工具被一步一步推到司法决策的中心位置的。公众对数字的需求，不是一个天生就有的社会学常量；它是在惩罚系统失去肉体可见性之后，在社会对司法系统的持续不信任与改革压力的反复互动中，逐步被生成、被强化、并最终被自然化为“理所当然”的一种认知习惯。

当这种习惯长成之后，“不可接纳者”这个极的填充物就发生了静默的更替。过去，这个位置是由道德话语（“他是邪恶的”）、法律论证（“他的行为符合重罪构成要件”）和传记叙事（“他从小就是坏种”）联合填充的。今天，在最依赖精算评估的那些司法管辖区，这个极越来越多地由数字痕迹来填充：一个人不是因为他内心被看透而变成不可接纳者，而是因为他的分数把他甩进了一个只能不断产出“危险信号”的生存轨道，然后这个轨道里的每一步都反过来硬化他的这个位置。他的“不可接纳”不是起点，而是终点——而且这个终点是由骨头为

他提前锁定的。

一位十九世纪监狱史学者菲利普·普里斯特利在其对维多利亚时代囚犯传记的研究中，曾捕捉到一种在数字时代之前就隐约可见的结构惯性。他发现，当时的监狱管理者时常困惑于一件事：他们把犯人关进单独监禁室，意在令其忏悔，但隔离往往只让犯人变得更沉默、更阴郁、更与社会疏离——也就是更“像”他们本来被指控的那个恶人。普里斯特利勾勒了这样一种机制：监狱系统倾向于把囚犯塑造成它预先假定他们就是的那种存在，然后把这一结果当作系统最初判断之正确的证据。[5]那个时候还没有COMPAS，没有精算评估，没有任何我们今天谈论的数字骨头。但这个机制，今天在大规模地、低成本地、自动化地运行着：先把人预设为一个类型，然后确保他回应这个预设的每一条路都通向这个类型的加深，最后把他被加深的结果归档为这个类型天然存在的证据。

当代社会为什么持续需要这个极？不是因为有个抽象的“社会需求”永远在那里等着被满足。而是因为每一次公众对司法系统的信任危机、每一次媒体对犯罪的集中报道、每一次选举周期里“打击犯罪”的立法承诺中，这个由数字痕迹标记的“不可接纳者”的位置都被再次激活、再次确认、再次填充。它的功能远远超出了“被惩罚”——它的存在本身，就为剩余的社会成员提供了一种可以随时被调用的反向确认。这根骨头每锁定一个人，就为所有人提供了一份可以被数字支撑的安全感叙事。被骨头挤出的这个极，因此不是惩罚系统的副产品；它是现代司法合法性再生产过程中的一种结构性必需品。一个越是将自己理解为拥抱数据客观性的理性司法秩序，就越是需要这种由非人化数字痕迹所标记的“不可接纳者”来填充那些制度自身的合法性焦虑无法自行消化的空白。

九、结论：拔骨，而不是修尺

本文试图提出的核心命名，可以收束为一句极简的话：累犯风险评分，不是一把被用偏的尺，而是一根长在社会日常决断里的数字骨头。这根骨头靠三样东西活下去——决断者卸下判断重负的渴望，被评分者用痕迹换取制度呼吸的唯一出路，社会旁观者借数字对抗不安全感的持续需要。这三者都不是存留在个体心理层面的“错误认知”，而是被镶嵌在制度压力、组织惯性与合法性再生机制中的结构性的依赖。要拆掉它，只靠证明它“测不准”是不够的——那只是在修尺，不是在拔骨。

拔骨要面对的真实问题，不是“没有评分怎么办”，而是：当社会在某一时刻不能再把决断外包给那串数字痕迹时，它还能不能耐受判断本身的颗粒感——那些说不清的地方、那些必须由某一个人签下自己名字的伦理风险？布克案之后的联邦量刑实践给出了一个初步的、并不振奋的答案：当骨头被从强制性的框架中松绑时，决断者并不会自动“学会自己站立”——许多人会去寻找同一根骨头的柔性版本，继续靠着它。而从“依靠骨头的柔性版本”迈到“用肌肉自己站”，需要一个制度环境——在这个环境里，案量压力被控制在允许法官充分说理的范围，上诉审查不以“偏离标准化的数字框架”为主要的改判理由，而公众与媒体的合法性感知能够被判决书里对个体处境的充分叙述所满足，不只依赖一个数字的简短断言。如果这些条件不改变，拔掉骨头的唯一结果，就是另一根骨头长出来。

更诚实地说：即便社会真的开始了“学习用肌肉站立”，也没有一个“学会之后”的黄金状态在终点等着。用肌肉站立，不是一个一次性习得的技能，而是一个需要反复承受的、更容易疲劳、更容易摔倒、且永远不会像靠骨头那样省力的常态。拔骨不是通向一个更纯真的决断状态，而是通向一个需要不断在矛盾中重新配置决断责任的过程——每一次配置都可能失败，每一次失败都可能诱使骨头重新长出。

证伪条件。如果在某个辖区内，累犯风险评分被彻底从剥夺性决策链中移除（量刑加重、假释拒绝、监管级别升高），而只保留在支持性资源分配端（心理辅导、职业培训、住房帮扶），结果出现了以下事实——被释放者的再犯率在五年内显著上升，并且公众对该司法系统的“安全性”信任大幅降低，以至于出现大规模立法动议要求恢复评分在剥夺端的权重——那么，本文的诊断就被事实驳倒了。因为这表明社会的日常决断并未寄生在骨头上，骨头被拿掉后，社会自己站得很好，且更安全。但倘若，在同一试验中，剥夺性链条被切断后，再犯率无明显上升，而决断者们——法官、假释官——却持续在寻找下一个可以帮他们说话的“准数字替代物”，那就说明，被拔掉的从来不是一把尺，而就是那根骨头本身——而我们或许才刚开始学习，不用骨头，用自己的肌肉去站立。

注释

[1] “数字骨头”这一隐喻并非严格的生物学类比，而是社会存在论意义上的一种构型：将累犯风险评分系统视为嵌入社会身体日常决断中的支撑性器官，其首要功能不是提供精确测量，而是为决断者提供可被共享、引用与免责的行动支点。

[2] “痕迹存在论依赖”是本文为诊断评分系统不可替代性而提出的核心概念。它指向这样一种社会机制：个体的数字化痕迹（分数、等级、标签）在制度环节之间流转，被各决断者当作该个体的“所是”本身来操作，而非当作辅助参考信息；社会决断链已结构化地附着在这些痕迹上，以至于一旦移除痕迹，决断者将无法在日常操作中维持其判断功能。

[3] 参见Bernard E. Harcourt, *Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age* (Chicago: University of Chicago Press, 2007)。哈考特在该书中对精算方法通过改变执法分配内生地制造犯罪分布这一机制进行了系统论证。

[4] 关于布克案及其后果的实证研究，参见美国量刑委员会的年度报告及相关学术文献中关于布克案后“指南遵从率”的持续追踪。本文对布克案后司法实践的描述综合了该领域内已发表的实证研究的基本共识。

[5] 参见Philip Priestley, *Victorian Prison Lives: English Prison Biography, 1830–1914* (London: Methuen, 1985)。普里斯特利在该书中通过对19世纪囚犯自传、日记与监狱档案的系统分析，揭示了监狱系统倾向于将囚犯塑造成其自身预设形象的机制。

参考文献

- Bourdieu, Pierre. 1980. *Le sens pratique*. Paris: Les Éditions de Minuit.
- Esposito, Roberto. 2002. *Immunitas: Protezione e negazione della vita*. Torino: Einaudi.
- Foucault, Michel. 1975. *Surveiller et punir: Naissance de la prison*. Paris: Gallimard.
- Harcourt, Bernard E. 2007. *Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age*. Chicago: University of Chicago Press.
- Priestley, Philip. 1985. *Victorian Prison Lives: English Prison Biography, 1830–1914*. London: Methuen.

深化增补：被放跑的最近邻、操作化与证伪

以下四节为编辑增补，与作者正文分开标注。

增补一：可证伪的检验设计

本文的 D 与 I 都强，短板在证伪：既有论证说明了评分为何拆不掉，却未说清什么观察会让这个说明失败。核心命题可检验化为一组相反预测。支点论预测：在决断者不承担个人责任的制度环境中（例如判决由合议庭集体署名且无个人问责记录），评分的采用率与依赖度应显著低于个人问责环境，即使两地的统计准确度相同。测量论预测：采用率主要随准确度变化，与问责结构无关。若跨法域比较显示采用率与准确度相关而与问责结构无关，本文即被证伪。第二条独立路径：在准确度被公开证伪之后（例如某评分工具的偏差被权威研究确认），若采用率随之显著下降，测量论成立；若采用率不变或上升，支点论获得支持——这一自然实验在近年已多次发生，可直接回溯检验。

增补二：痕迹存在论依赖的可测形式

该机制的可测形式是替代率：在一次制度操作中，关于当事人的数字痕迹被援引的次数，相对于关于其完整处境的陈述被援引的次数。比值随时间上升，即为依赖加深。第二项是拔除成本：当评分输出不可用时，同一决断所需的时间与所要求的说明长度的增幅——若增幅极大，说明该输出已承担支点功能而非信息功能。两项都从制度文书取值，不需要访谈决断者的自我认知，因此可绕开“他会说自己只是参考”这一自陈偏差。

增补三：与算法治理批评文献的边界

算法治理批评已充分论证黑箱、偏差与自动化不平等。本文与之最易混淆，分野在批评对象：既有文献批评算法做了什么（错误分类、放大不平等、不可解释）；本文追问算法为什么拆不掉，答案不在算法的属性里，而在决断者与判断重负之间的张力。这个差别有实践后果：若病在算法属性，改进算法（去偏、可解释、可申诉）就是出路；若病在支点需求，改进算法只会造出一根更结实的骨头。本文因此预测一条既有文献不预测的现象——技术上更公正的评分工具，其扩张速度应快于被批评的旧工具。

增补四：与作者已发表工作的分野

作者已发表《恶法的腿：制度惯性如何让被道德遗弃的条文站住》，结构与本文高度相似——都解释某物在理由被赢下之后为何仍站立。分野在支撑物的位置：那篇的支撑来自制度惯性，是系统侧的延续力；本文的支撑来自决断者的个人处境，是使用者侧的需求。这也决定了拔除路径不同：惯性要靠制度换代，支点要靠替换决断者赖以站立的东西——若只撤掉评分而不给出替代支点，本文预测它会以另一种形态回来。这条预测是本文与前篇之间最有价值的差别，宜在文中明确。

编辑增补所依据的核验文献

- Bernard E. Harcourt, *Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age*, University of Chicago Press, 2007.
- Frank Pasquale, *The Black Box Society*, Harvard University Press, 2015.

Cathy O'Neil, *Weapons of Math Destruction*, Crown, 2016.

Pierre Bourdieu, *Le sens pratique*, Éditions de Minuit, 1980.

Roberto Esposito, *Immunitas: Protezione e negazione della vita*, Einaudi, 2002.