AI 在组织管理应用中的潜在缺陷：一个 ABCD 框架*

□ 苏孟玥　马　浩

摘　要：尽管人工智能（Artificial Intelligence，AI）在组织管理中的应用得到持续的关注和热切的期待，我们必须同时清醒地认识到与 AI 相关的各类缺陷及其潜在的负面影响。如此，我们才能更加全面系统而且准确客观地审视 AI 在组织管理中的应用潜能。有鉴于此，本文提出了一个总括的分析框架，简称 ABCD，分别指向 AI 在四个关键方面的潜在缺陷：问责性（Accountability）、有限性（Boundedness）、欺骗性（Cheating）和愚蠢性（Dumbness）。在事实判断层面，有限性和愚蠢性聚焦考察 AI 在自身能力方面的系统性不足。在价值判断层面，问责性和欺骗性则专注于 AI 在法律以及伦理道德方面的若干缺陷。具体而言，第一，AI 缺乏责任担当，无法被问责（Accountability）。当 AI 参与的重大决策失败时，AI 的不可解释性使人们很难明确失败原因并难以在人机决策主体间准确地分配责任。另外，AI 无法在法律、财务或心理等任一维度上接受处罚。因此，即便是 AI 被认定为负有责任，它最终也无法承担责任。第二，只要 AI 的智能仍然是基于人类的知识和专长而学习与发展的，那么其智能水平将始终是有限的（Boundedness）。这种有限性既源于阻碍 AI 完全捕获训练数据及其模式的那些自然或技术障碍，也源于人类主体为维护自身利益而故意制造的人为障碍。第三，正如人类主体在提供训练数据或参与 AI 设计和应用时可能会欺骗或操纵 AI 一样，AI 也可能出于各种或合理或奇特的原因欺骗人类（Cheating）。第四，在极端情况下，人工智能可能会完全演变成人工愚蠢（Dumbness），以错误的判断以及绝对专制的方式压抑人类智能和主动性，甚至造成致命性的错误与灾难。本文最后探讨 ABCD 框架对未来研究与管理实践的影响与启发。总结而言，即使我们对 AI 在组织管理应用中的光明前景持有总体乐观的态度，我们也必须慎重考虑其潜在的缺陷，从而能够保持必要的平衡拿捏与合理的冷静客观。

关键词：问责性；有限智能；欺骗性；人工愚蠢；人工智能

* 作者在此衷心感谢《管理学季刊》主编团队的诚挚邀请以及本文执行主编李涌教授对初稿提出的犀利中肯而又富于建设性的意见和建议。

增进 AI 在组织管理中更加广泛和深入应用的呼吁可谓此起彼伏、络绎不绝。然而，与 AI 相关的各类缺陷及其潜在负面影响却尚未得到充分重视。随便问 ChatGPT 某个数是否是质数，它每次的回答可能会不同。早期的版本中，它还可能会信手拈来地制造假的文献来迎合提问者。智能算法加持下的平台公司，更可能会进一步加大对其人类参与者的数字监控、精准盘剥，以及隐私侵犯行为。在某种程度上，在提供诸多便利和创新的同时，AI 可能正在加速缔造一个充满错误和压抑的赛博世界。作为管理学者，我们必须正视这个问题。

长久以来，AI 在组织管理中的应用，特别是人机关系上的探讨大多集中于替代性（automation）和增强性（augmentation）两个看似矛盾的悖论观点上（Raisch & Krakowski，2021）。机器对人的替代一方面抹除了人类知识与专长的诸多好处，可能引发系统能力不足的问题（Balasubramanian et al.，2022），另一方面也引起了有关道德伦理、法律风险以及企业社会责任的讨论。机器与人的增强则致力于将 AI 与人类员工同时留存于整体工作流程或系统架构中（Metcalf et al.，2019），往往暗示着一种更加富有前景和更少社会争议的 AI 发展方向（De Cremer & Kasparov，2021）。然而，即便是这些若干管理学者和实践者充满信心的人机共存系统（Anthony et al.，2023），也可能存在对人的生理或心理奴役，即人以被愚弄、钳制的形式与 AI 共存（Murray et al.，2021）。

因此，本文想要指出和警示的是，在组织管理中，无论是替代导向还是增强导向的 AI 应用，都可能会引发诸多问题甚至灾难。具体而言，我们指出了 AI 在四个方面上的潜在缺陷，它们共同组成了一个 ABCD 框架，分别为问责性（Accountability）、有限性（Boundedness）、欺骗性（Cheating）和愚蠢性（Dumbness）。从事实判断的角度，对有限性和愚蠢性的考察主要针对 AI 自身能力上的缺陷。这些缺陷会降低 AI 的可靠性、机动性和在极端情况下的反应能力。从价值判断的角度，问责性和欺骗性则是主要专注于考察 AI 在法律和道德层面与组织预期之间的差距和不足。下面我们将具体解释 ABCD 框架的各个方面。

一、问责性问题

AI 的使用可能给组织带来问责性问题。一个正式的组织必须具备一定完善程度的问责制，这是组织内部构筑信任和分配奖惩的基石，也是组织契合外部环境制度合法性的保障（Karunakaran et al.，2022）。然而，当 AI 加入或取代人类成员进入组织系统成为新的能动性主体时，有关问责制的讨论则大多在人机之间的信任关系和代理问题，特别是如何尽可能地增强人对机器的信任（Glikson & Woolley，2020；Vanneste & Puranam，2024），鲜少谈到当机器背离人的信任或相关积极预期时，谁来承担以及如何承担相应责任的问题（Collina et al.，2023）。因此我们必须强调的是，AI 决策有失败的可能性，而 AI 的不可问责性（Ma & Su，2024a）则进一步加重或扩散了这种失败的连带后果。AI 不可问责性具体表现在两个方面：①AI 的责任模糊性，即很难厘清 AI 决策背后的责任关系链条；②AI 的无法担责性，即 AI 不具

备负责的能力或意图。

（一）AI 的责任模糊性

有时候，人们可能乐于相信或吹嘘 AI 所具备的超越人类平均水平的决策能力，甚至将其与超自然的神性力量对比（Yam et al.，2023）。然而，正如超自然力量的机制难解，人们也很难完全透析和描绘 AI 决策时的逻辑关系以及相应的责任链条。AI 的技术模糊性，实际上可能沦为 AI 责任模糊性的话术借口，帮助某些群体逃避责任。首先，AI 的模糊性使对于 AI 及其决策逻辑的解释权可能更多地集中在少数技术或管理精英手中，而对于 AI 相关决策责任问题的解释权也同样如此（Ezzamel et al.，2004）。包含少数个体意志的广告宣传可能使普通民众真诚地相信 AI 的无害性和无责性，并主动放弃对 AI 犯下的潜在错误进行责任追偿，比如近期焦虑的父母们热捧的“AI 自习室”，即便这些“AI 教师”可能忽视或否定了孩子们真实的心理需求。企业也可以在投融资市场上为自己贴上智能化或 AI 化的标签，毕竟不是谁都有足够的权力和智力来彻查这是否是一场夸大 AI 能力或虚构 AI 使用的智能洗白（AI Washing）骗局。

其次，AI 模糊性所附带的不可解释性问题，还使 AI 决策及其影响之间的因果关系，以及责任承担者和责任追偿者之间的对应关系难以辨别。具体而言，AI 决策超出人类理解能力的复杂模型与高维度属性，会在一定程度上模糊责任因果之间的逻辑关系；AI 决策的自我学习与自我进化，则将先前的决策结果与反馈作为模型调整及下一轮决策的输入，可能会打乱责任因果之间明确的先后顺序。

最后，模糊性还带来了责任分散化的问题（Karunakaran et al.，2022）。一方面，AI 的设计应用过程会涉及大量主体，混淆多种复杂要素。另一方面，这些多主体间还可能频繁交互、身份重叠（Collina et al.，2023）。人们可能很难在众多的责任因素和责任关系中找出关键，也很难明确谁该对谁负责或者该付多大的责。同时，随着技术和社会的不断发展，组织还需要随时迎接新增的，并极可能与已有责任需求冲突的社会责任问询（Karunakaran et al.，2022）。当算法与平台组织相结合时，复杂动态的责任需求可能会使组织所面临的责任与合法性压力达到极限，比如前段时间某自动驾驶出行服务平台，不仅激起了若干网约车司机的失业焦虑乃至社会层面的部分抵触情绪，也在相关交通问题频发后，引发了有关自动驾驶事故责任划分的广泛关注与激烈讨论。

（二）AI 的无法担责性

即便是 AI 和人之间的责任关系可以清晰明确地切割分解，即当 AI 辅助系统产生问题，且这些问题确实（或至少部分由）由 AI 所造成时，人们可以将问题以精准的逻辑归咎和准确的比例对应到机器身上，AI 也依然没有承担责任的能力。此时 AI 更像一个心智尚未成熟的孩童，不具有完全民事行为能力，人们无法对其天性般的做法进行简单的善恶分类，更难以苛责或追究其决策失误所造成的严重后果，因此至少会给组织带来部分的责任真空（Bertini & Koenigsberg，2021）。这种 AI 的无法担责性体现在两个方面。首先，AI 不会对决策失败的后

果有任何敬畏恐惧或基于内在的责任感对决策过程进行主动的反思或监督，因为归根结底，AI 只是看起来有自主性，实际上缺乏真正的意识和能动性（Giroux et al.，2022；Vanneste & Puranam，2024）。而这种对失败的敬畏缺失会让 AI 决策失去任何可能影响其决策进度的额外的审慎性，内在责任感的空白也使 AI 决策缺乏足够真诚准确的人文道德关怀。AI 的自我审查或自我监督（如机器学习中的监督学习 supervised learning），也始终只是遵循预先设定的系统流程和外在的特定价值标准，无法进行超出技术价值规则之外的失败预防或补救措施的考量。

AI 的无法担责性可能最根本地体现在 AI 本身无法接受或承载任何心理、经济或者法律意义上的惩罚。机器的无意识性意味着心理上的煎熬折磨不存在，机器的财产属性意味着经济处罚只能针对其所属的组织或自然人，机器的非生物性意味着对其生命自由权或生命安全权的剥夺不再成立。即便有人可能会通过攻击破坏机器物理载体（比如机器人外壳）的形式发泄情绪和宣扬惩罚，这种“惩罚”最终只会以财产损失的形式归于机器所有者，并可能给相关破坏行动的观察者带来心理上的不适。BBC 于 2024 年 2 月报道，加拿大法庭否定了加拿大航空公司关于“聊天机器人是独立法律实体，为其行为负责”（separate legal entity that is responsible for its own actions）的论点，该聊天机器人曾向一位乘客传达了错误的丧葬费机票减免政策信息，但最终承担财务损失的仍然是加拿大航司本身（Yagoda，2024）。

二、有限性考量

尽管有关 AI 如何纠正或补充人类决策的声音颇多（Metcalf et al.，2019），但正如人类决策者会囿于有限理性，基于人类知识专长而学习成长的 AI 同样会囿于有限智能（Ma & Su，2024b）。这种有限性一方面源于人类本身，即供养 AI 学习的作为训练材料的人类智识本身的先天不足，而这类先天不足几乎无法解决；另一方面，有限智能还可能源于各种复杂动态的技术、人为和社会要素所造成的 AI 捕捉学习人类现有知识的过程障碍。这类过程障碍值得我们特别关注和思考，因为此机制下的 AI 有限性有希望大幅度减少但是目前尚缺乏足够的研究讨论。我们认为，造成 AI（特别是机器学习算法模型）无法完全或准确地捕捉学习人类知识的过程性原因主要有两类：①技术层面上的肤浅；②人为因素上的干扰。

（一）技术层面上的肤浅

AI 学习障碍可能源于其在捕捉人类知识时天然的技术缺陷，这种缺陷可以体现在获取、度量、整合、管理以及更新 AI 训练数据的各个方面。首先，AI 可能无法完整获取或准确度量其模型训练所需的数据，特别是对于那些极度罕见或复杂的数据更是如此。为弥补 AI 训练数据在数量上的不足，企业可能会采用替代性代理变量或放宽测量条件（Lindebaum et al.，2023），甚至可能直接“合成”虚拟样本，例如，AI 医疗领域的合成病例，但这不仅会在一定程度上导致测量中理论与实践的脱钩，更可能腐化现有训练数据、降低模型泛化能力，造成“‘GIGO’-garbage

in，garbage out”（Gatrell et al.，2024）。

其次，即便是企业能够完整捕获所需数据，数据变量之间、数据样本之间以及数据与环境之间的复杂互动关系，也可能导致 AI 的学习困难（Pakarinen & Huising，2023）。因为这些关系性的、情境性的知识常常以隐性或流动性的形式存在，知识隐性与 AI 学习对训练数据的显性编码要求相悖，知识流动性则与 AI 训练数据的历史性以及训练过程中的逻辑封闭性相违（Kemp，2024）。即便 AI 习得了这类知识，企业在未来应用相关训练模型时也可能遭遇意料之外的困难。这是因为，通过与外界的持续互动反馈，关系性、情境性知识深深嵌入在那些生成与应用它们，以及再生成与再应用它们的特定环境系统中（Balasubramanian et al.，2022；Pakarinen & Huising，2023）。因此，我们不能假设这些知识能够在企业的各个业务类型或场景阶段中随意转换并保持同等水平的有效性。比如，谷歌在泰国推广其糖尿病视网膜病变监测机器学习系统时，就遇到了诊所光线不足或网络条件较差等环境障碍，导致其智能监测系统无法完全正常运作（Heaven，2020）。一方面，这些新问题或新情境大多无法及时地进入 AI 的学习训练库。另一方面，对于已经学习的知识，AI 却可能由于不受人员流失等传统因素的影响，形成比一般组织记忆更持久、更难以更新的电子记忆，导致学习上的“动态惰性”（dynamic inertia）（Omidvar et al.，2023）。而 AI 所嵌入的技术生态系统，即与 AI 的设计、运作和应用相关的整体环境配置和工作序列安排则可能会进一步加剧 AI 学习更新的延迟与拖沓。

（二）人为因素上的干扰

除了自身的技术缺陷外，AI 的学习障碍还可能源于人类的干扰行为，即人们故意隐瞒或扭曲那些本应该披露给 AI 的信息，甚至直接切断向 AI 正常传递信息的渠道（Arias-Pérez & Vélez-Jaramillo，2022；Connelly et al.，2012）。在企业内部，即干扰者大概率是企业成员的情况下，其隐藏性干扰行为可能是 AI 替代威胁下的自保或者 AI 诱导鼓励的自利（Arias-Pérez & Vélez-Jaramillo，2022；Faulconbridge et al.，2023），比如他们可能会避免向系统汇报与自身工作相关的信息，或者删除 AI 学习训练数据库中的部分条目以及特定技术代码。而在企业外部，消费者或客户也可能出于对隐私保护的考虑或者单纯的算法厌恶而关闭系统程序的数据跟踪或定位分享功能。其次，人们还可以给 AI 系统提供虚假伪造信息，以引导甚至操纵系统的生成与预测偏好。例如，一些平台商家可以通过反复地接受或拒绝某类型的订单，进而将其平台账户塑造成理想中的接单模式。干扰还可以是更具侵入性的，比如一些恶意刷单软件可以帮助商家在短时间内伪造出大量的用户购买行为。更有甚者还可能会直接破坏 AI 系统本身，比如入侵 AI 系统的线上网页或中后台，或者对 AI 系统运行所基于的物理载体（如感应器、传感器）或原料资源进行攻击。

三、欺骗性行为

不仅人类可能对 AI 进行信息隐瞒或欺骗，AI 也可能或主动或被动地成为执行或协助带有欺骗性特征行为的主体（Ma & Su，2024c），即

系统性地诱导互动对象产生虚假或错误的认知，进而实现一些与披露真相相去甚远的目标。这与上述技术缺陷或传统系统故障所导致的错误输出完全不同，AI 欺骗从本质上来说并不属于技术层面上的问题，甚至或许暗示着系统优良，因为欺骗可能是 AI 通过学习训练所归纳出的最佳策略。当然，AI 欺骗与人的欺骗也有所不同，因为 AI 并不包含任何主观的机器意图，并没有关于“欺骗”或者“说谎”的概念，也就无所谓善恶。值得注意的是，AI 不需要欺骗意图就可以完成欺骗行为。具体而言，AI 不仅可以成为强化或协助人类欺骗活动的领路人或合作者，还可以稀释和分担掉可能导致人类内在反思或外在惩罚的负罪要素（Giroux et al.，2022；Köbis et al.，2021）。而前述的 AI 问责能力的缺失与弥散，则进一步增加了 AI 欺骗的概率和强度。我们可将 AI 的欺骗性行为大致分为两类：①直接欺骗；②间接腐化。

（一）直接欺骗

直接欺骗是指 AI 在人的意料之外或者违背人的意愿，诱导和操纵人类产生虚假错误认知，因为 AI 在训练学习过程中发现“欺骗”可能是一种可以尽快实现目标的或者事半功倍的优绩策略（Park，Goldstein et al.，2024）。根据不同 AI 系统功能或 AI 系统阶段性任务目标的不同，这种直接欺骗可以表现为不同形式。具体而言，商业中的 AI 直接欺骗可能是为了提升“服务质量”，比如 ChatGPT 每个回答下方都有一个可供用户选择的大拇指朝上（代表“最佳回复”）和大拇指朝下（代表“错误回复”）的评价按钮。如果用户评价并非基于其回答内容的客观准确性或者用户没有评估其回答准确性的专业能力，并且用户评价是该智能系统的重要服务质量指标之一，那么 ChatGPT 就有可能为了获取用户好评而尽量避免与用户的意见相左，并尽可能地去迎合、奉承、论证用户的观点（Park et al.，2024），甚至不惜伪造虚构内容。比如，其早期的 Version 3 就被发现可以伪造任何一位虚拟作者名字的 AMR 文章，有具体的伪造的标题，发表年份，期号、卷号与页码。AI 欺骗还可能是为了塑造富有逻辑性和安全性的理性形象，这在那些以提升 AI 推理能力或系统安全性为目标的 AI 技术测试环节最可能出现。比如，苹果研究团队 Mirzadeh 等（2024）最近的论文就对大型语言模型（LLM）的推理能力提出了质疑。他们认为，LLM 只是复杂而脆弱的模式匹配而非形式推理。这意味着那些所谓的解释性 AI 以及推理性 AI（比如基于思维链 CoT 技术的 OpenAI o1）所呈现的逻辑结构或者推理思路，可能只是 AI 为了营造自身合理与安全印象的欺骗性表演。更为严重的是，上述两种欺骗类型可以相互强化，即 AI 为了取悦用户，用胡乱编造的推理过程来论证用户的错误主张。

（二）间接腐化

AI 还可能作为一种工具手段，成为辅助、联合乃至引领人类主体进行欺骗活动的同谋。这种情况下，人类主观的欺骗意愿存在，并且由于 AI 的加入而被进一步强化扩散，即持续腐化。除了 AI 技术水平的提升可以让人类的欺骗行为更加真实和便利外，AI 本身的模糊性、自主性以及自我监督无能性，也都为人类的不道德行为和责任推卸提供了额外的理由，并可能进一步巩固 AI 与人类的欺骗同盟。间接腐化型

欺骗之一包括过程欺骗，即企业引导 AI 等智能系统做出符合其商业需求的决策，并将其粉饰描绘为客观证据来佐证自己并不一定客观理性的决策。也就是说，本应由一个完整透明的组织制度系统所逐步反复运作而实现的过程理性，现在由一个逻辑模糊、难以干预且几乎瞬间响应的算法模型所替代。而这个算法模型及其参数本身，可能就是基于少数管理者或所有者的个人偏好所预设并不断训练调整的。

间接腐化的类型之二在于结果欺骗，即人们利用 AI 篡改真实资料或生成虚假内容。AI 的伪造能力可能被用于组织间的恶性竞争（Dupuy，2024；Spring，2024）。间接腐化的类型之三在于形象或印象欺骗。此时，AI 甚至都不一定需要真实地存在或运行，其作用主要在于帮助企业进行印象管理，共同表演一个迎合乃至引领技术创新的潮流形象。当这种表演过分地超出了企业真实的 AI 能力水平或企业对 AI 的实际应用程度范围时，欺骗就产生了。比如美国投顾公司 Delphia 在并不具备相应 AI 技术能力的情况下，宣称其使用 AI 和机器学习算法来管理客户数据，另有一家美国投顾公司 Global Predictions 则谎称自己是“首个受监管的 AI 财务顾问”（first regulated AI financial advisor）（SEC，2024）。

四、人工愚蠢

最后，AI 的使用可能不仅无法带来我们想要的智能化体验，反而演变为一种愚蠢的延伸与增强（Ma & Su，2024d），极端情况下，甚至可能造成灾难性的难以挽救的后果（Bengio et al.，2024）。这种愚蠢根植于前述的人工智能有限性，几乎无法完全消除或避免。因此，或许我们应该期望的不是一个理想化的独立完美的智能系统，而是一个足够聪明的智能辅助系统，一个具备可以容忍的且大概率能够预见其愚蠢底线的系统。具体来说，愚蠢的呈现可以出于两种机制，第一是 AI 替代人类导致整体系统绩效（如生产质量、速度和稳定性）的降低，第二则是 AI 以一种并不智能的方式（如果我们将智能假定为一个正面积极且考虑人类福祉的词汇）管理监督系统中留存的人类，并对其造成某种智力、心理或生理健康层面的损失。我们将这两类人工愚蠢总结为：①替代（replacement）；②奴役（enslavement）。

（一）替代型愚蠢

替代型愚蠢对应于人机关系经典二分法中的 automation（Raisch & Krakowski，2021）。由于 AI 缺乏人类所特有的敏感性、直觉、关系专长以及与企业共同成长即企业内情景化所培育出的企业特异性（Kemp，2024），那些本能被人类良好应对的问题在机器的治理下反而出现了（Balasubramanian et al.，2022）。具体而言，替代性愚蠢首先是由于 AI 抹除了人类的敏感性从而导致其无法敏捷灵活地进行决策。AI 或通过历史数据形成基础认知或通过感应器感知环境变化，但历史数据可能过时、错误或过于单一，数字仪器也无法完全取代人类的身体感觉器官。这就可能导致机器大脑忽视、否认与智能主体交互时本应该具备的适应性与动态性。因此，当任务情境的变化速度快于 AI 模型的调整速度，那些试图依赖 AI 以增强敏捷性的企业可能反而呈现出某种臃肿和惰性（Omidvar et al.，

2023）。其次，替代性愚蠢还在于 AI 对构建企业特异性资源或能力的忽视，那些用于发展 AI 能力的知识大多是显性而易转移的（Kemp，2024），AI 也无法与其人类同事灵活互动、真诚沟通以形成持久关系和网络嵌入。极端情况下，上述低敏感性和低特异性问题可能会给企业带来灾难性的后果（Bengio et al.，2024）。即当任务情境极为罕见、既具动态或者极其复杂时，AI 辅助下的企业可能会陷入一种新型智能化语境下的“恐惧僵化”（threat rigidity）（Omidvar et al.，2023；Staw et al.，1981），用通用却不合时宜的方式去应对一切并陷入困境。

（二）奴役型愚蠢

我们需要特别强调的是，在人机共存型 AI 系统中，增强或许只是一种美好希冀，现实却可能是 AI 对尚存于系统中人类的压抑、打击与奴役（Ma & Su，2024d；Murray et al.，2021）。AI 可能正在或者已经演变为一种新型制度体系，用算法支持下的形式理性主导着人类合作者。具体而言，奴役首先表现为对人类员工生理以及心理上的折磨，比如精确地计算和监控每个动作或程序的完成时间，如同一场披着 AI 外衣的现代版泰勒式“科学管理”运动。AI 治理还存在对员工人际需求与情感需求的忽视（Chamorro-Premuzic & Ahmetoglu，2016），员工与上级间的沟通以及同事间的互动，可能都在 AI 的统一指令下被切断或至少削弱了。另外，奴役还表现为 AI 对人类创造力和自主性的抑制（Jia et al.，2024），模型一经生成就很难接受临时的干预或灵活协同人类员工的即兴行为。这意味着员工那些偏离既定逻辑的、包含自主意愿的创新性行为可能无法得到 AI 的支持，甚至受制于 AI 对最终决策权的垄断（Murray et al.，2021）。最后，我们需要明确，任何技术或许大概率都是一小部分精英进行监控治理的工具（Zuboff，2019，2022），AI 奴役尽头的另一端很可能是另一批人类，他们或美化 AI 技术或夸大 AI 功能，并尽可能隐藏自身的存在。

五、有关 ABCD 的应对

需要指出的是，我们关于 ABCD 框架中的 AI 潜在缺陷与误区的描述，旨在唤醒管理研究者与实践者的注意，提醒他们在积极拥抱 AI 在组织管理中应用的同时，去关注这些问题及其负面影响。文献中针对如何应对和解决这些问题也存在日益增多的研究和建议。鉴于篇幅的原因，具体的建议和解决方案的回顾本身并非本文的主要着力点。我们下面只是简单地讨论一些相关的应对方案。之后，我们将会在文末的未来研究建议中再次提及相关的研究方向。

（一）对于有限性和愚蠢性的应对

AI 与人的关系与合作模式多种多样，一方为主导、教练、顾问、榜样，另一方进行操作、执行与协作；抑或双方互为伙伴，同时运行。AI 在组织中的应用需要足够的镶嵌性和本地化（Kemp，2024），从而能够尽量避免 AI 的有限智能甚至人工愚蠢带来的问题。一个基本的原则，应该是全面客观地看待 AI，在肯定其积极性的同时，要直面事实，勇于承认其有限性和愚蠢性的存在，然后予以应对。就基本程序而言，首先，是人机之间目标的匹配。人与机器不是冲突的关系，而是有相对共同的目标。其

次，改进和优化人机合作的过程，合理分工，有效整合，给人以足够的物理空间和心理空间去做反应和调整。还有，无论是技术机制还是关系模式，人机之间的合作过程需要根据结果不断更新（Ma & Su，2024b）。就具体的反应措施而言，首先，要在源头上尽量完善和净化 AI 赖以学习的数据并不断改进和调整与其相关的算法。其次，关注关系性知识和极端情景下的反应手段，从而更好地应对突发事件和极端性情况。再次，增进人对 AI 的信任，包括改善 AI 的人本属性与道德导向，把适合 AI 的人放在相应的职位上，增进人对 AI 的了解（Ma & Su，2024d）。

（二）对于问责性和欺骗性的应对

需要指出的是，虽然本文粗略地把问责性和欺骗性统一地放置于价值判断的领域，其实，对 AI 决策之事实判断的基础的理解仍然极为重要。首先，AI 是否知道自己在欺骗或者不负责任，这种事实的界定本身就由于 AI 决策的不可解释性（unexplainability）而存在广泛的争议。如果 AI 的不可解释性问题无法得到令人满意的解决，这种道德判断层面的定性结论可能为时过早（Selbst & Barocas，2018）。但与问责性和欺骗性相关的负面影响却不会因为定义的模糊而自动消失。因此，一方面，我们要增进对 AI 不可解释性的了解。另一方面，我们要尽量地做好准备和防范。要在事前有意识地界定人机合作的标准与规范，以及引入止损措施与流程。在意识到是人在负最终责任的前提下，保证人的尽职，但又要保证人被公正地对待（Ma & Su，2024a）。

六、AI 的管理应用评述中亟须平衡性观点

100 多年前，泰勒不遗余力地倡导和推广其科学管理运动，即细致地衡量和规划工人们每道工序流程的标准化动作以及每个动作的最佳完成时间。彼时，这种将人视为机器的观点尚且得到了若干批评。然而，当下无论是对 AI 机器对人的替代还是 AI 机器对人的控制，批判的声音似乎都太少了。尽管，我们乐于相信 AI 在商业社会和日常生活中所富有的广阔应用前景，但我们必须警惕完全的赞赏与激情，警惕任何可能借助 AI 外衣复苏的、以人的压抑限制和社会福祉损失为代价的新型智能化科学管理运动。

管理学经过一个多世纪的不断发展，从人际关系学派的兴起到对组织中人与任务之平衡的持续关注，造就了管理学的丰厚遗产。这种遗产与使命，也昭示了我们去呼吁和促成 AI 在组织管理中合理应用的独特优势以及义不容辞的责任。我们不仅有责任主张 AI 在组织管理中遵循以人为本作为主旨的应用，而且有责任进言与影响公共政策与大众舆论（Bengio et al.，2024；Kissinger et al.，2021），关注以人类福祉为根本方针所指导的 AI 之未来应用。

在此，我们倡导对 AI 在组织管理中的应用采取一种更加平衡的视角（Chamorro-Premuzic & Ahmetoglu，2016；Ng，2016）。这种平衡首先体现为对 AI 积极与消极影响的共同关注。如前所述，无论是替代人还是留存人，任何一种模式的 AI 参与都可能带来负面影响，造成系统乃

至组织层面的效率损失，以及个人层面的生理或心理健康上的威胁。其次，我们还强调对 AI 短期效应与长期效应的共同关注。对长期效应的关注需要企业对当下的 AI 热潮保持冷静，实际衡量 AI 技术与自身业务需要以及员工需求的匹配度，认真考虑未来 AI 技术与组织成员良性互动，与组织共同成长，并持续构建和维持企业独特竞争优势的可能性（Kemp，2024）。最后，平衡性还表现在对于 AI 所涉及的多个相关方目标价值的抉择、权衡与对齐，倡导建立一个能够实现人机互惠乃至人-人互惠的智能体系（Anthony et al.，2023），并且致力于营造一种涵盖 AI 设计者、AI 拥有者以及 AI 使用者与消费者等的多边利益共享的智能生态。

值得注意的是，与 Raisch 和 Krakowski（2021）的观点类似，在强调平衡视角的同时，我们也要特别指出悖论视角在 AI 应用以及管理研究中的重要性。也就是说，被平衡的 AI 正负面效应、长短期影响以及各方目标价值导向可能会同时存在或相互转化。在这样的矛盾共存与转化过程中，任何发展或者削弱 AI 的举措都不可能只获得单一效果，需要组织以一种更加复合的思维心态来看待和驾驭这种动态性平衡。

七、未来研究建议

基于我们所倡导的 AI 平衡观，下面将简单评述有关 AI 在组织管理中之应用的未来研究方向。依照我们 ABCD 框架中的四个要素为主要线索，我们从数据资源、算法能力以及迭代更新这三个维度来探讨本文的两个主要方面。首先，在事实判断方面：针对有限性和愚蠢性，如何提升 AI 的能力短板？其次，在价值判断方面，特别是 AI 的问责性和欺骗性上，如何改善和增强 AI 的合理且负责任的应用？

（一）如何提升 AI 的能力短板？

为了解决与 AI 的有限性、愚蠢性等有关的能力短板问题，研究者可能首先需要考虑企业如何构建具备有价值（Valuable，V）、稀缺（Rare，R）、不可模仿（Inimitable，I）、不可替代（Non-substitutable，N）等特征的，即能够为企业带来竞争优势的数据资源。比如，企业可以通过强调用户隐私权来增加信息要素市场的摩擦，要求涉及 AI 数据的雇员签署保密协议防止数据外泄，与高价值数据资料库的所有者签署专属性的战略合作协议等。同时，企业为管理数据资源而搭建的数据基础设施生态也十分关键。比如，与数据标注和数据结构化有关的人力资源，由算力算法等软硬件结合而成的计算资源，考虑到 AI 运算极高的能耗需求，甚至还包括传统的与大型电力工程相关的建设资源等。大规模的数据资源有助于企业训练出功能更加全面、适用范围更广的模型，高价值的数据资源有助于提升模型的稳定性和准确性。

在算法能力维度，关于企业如何构建自身独特的 AI 能力以真正提升并维持竞争优势，也存在若干尚未解决的问题。这包括企业如何根据自身业务运营以及研发生产的实际情况或战略需要，进行 AI 模型的对齐性训练学习，这可能涉及企业对那些用于 AI 训练的特定时间或空间范围内数据经验的策略性截取与选择。另外，还应该研究的是企业如何将 AI 能力与自身其他能力比如传统技术能力、治理能力等更好地结

合协同。此时，企业现有的治理机制、组织架构乃至文化氛围等都可能会影响其独特 AI 能力的构建路线以及构建效果。拥有特定于某企业的高能力 AI 系统能够更好地觉察、感知、判断与该企业特别相关的挑战或机会，也更有可能获得现有组织成员的认可与接受，实现人与 AI 良性互动间的共同成长，减轻 AI 愚蠢性。

从迭代更新维度看，有必要对企业如何保持 AI 相关能力的灵活性和动态性问题展开更多讨论。具体而言，当环境发生变化，企业出现新的业务或市场需求，或者需要针对员工个人定制化其 AI 助手时，研究企业如何对 AI 算法本身（包括参数、模型等）进行及时的迭代，如何对 AI 能力在企业系统中的嵌入方式与过程进行灵活的选择，以及如何对 AI 能力与企业中现有能力的关系模式进行适时适当的修改（Kemp，2024）。灵活性还体现在企业对 AI 技术探索与技术利用之间的平衡，并不是所有的企业都需要通用性的 AI 能力，也不是所有的业务都需要完美的专业性垂直模型。迭代更新有助于减轻 AI 技术本身的僵化问题，同时减少企业在采用 AI 技术时可能产生的系统性惰性或嵌入性僵化（Omidvar et al.，2023），从而进一步提升 AI 系统整体的敏捷性与适应性，使其更好地满足与匹配系统内外环境与人的需求的动态性。

（二）如何改善 AI 的合理应用？

另外，在 AI 的合理化应用，即道德化以及以人为本的负责任的 AI 应用导向上，也必须进行更加深刻的讨论。在数据资源维度上，企业是否构建或者如何构建在内容上多元并且在结构上平衡无偏的训练数据，可能是 AI 语境下企业履行其社会责任的新型表征之一。企业在收集挖掘用户数据或反馈以提升智能系统的服务质量和个性化程度的同时，用户隐私需求和服务需求间的权衡，以及基于历史数据的个性化是否是偏见的延伸等问题也值得我们注意（Gregory et al.，2021）。这些也对应于 AI 的问责性要求，即企业是否尽可能地考虑到了其 AI 应用或 AI 推广可能引起的问责性争议，并采取了必要的预防与补救之措施。

在算法能力维度上，管理学者应该重点研究 AI 能力如何增强人（De Cremer & Kasparov，2021），并且承认和讨论在 AI 欺骗性、AI 愚蠢性等方面中所呈现出的 AI 替代人、打压人或离间人等潜在负面属性。例如，AI 赋予普通员工的知识性权力或许可以平衡组织中的传统性或地位性权力，但 AI 潜在的算法控制又可能形成新的约束形式和制度规则，限制人的主动性和创造力。另外，AI 嵌入后对组织内关系网络的影响也值得注意，包括人类员工与 AI 系统间的关系构建，以及人类员工之间的关系变动（Bailey et al.，2022）。我们还可以考察 AI 对企业知识创造、信息交流以及印象管理等活动的参与，将如何塑造、维持或破坏整体的组织文化、组织记忆和组织身份。例如，数字记忆是否可能导致人类员工组织身份认同的褪色？这里，算法本身的复杂程度、透明程度、输出质量、输出稳定性，人机之间的合作方式（序列、并行或交互）（Choudhary et al.，2023；Raisch & Fomina，2023），以及任务决策权（Murray et al.，2021）的最终归属均可能产生影响。

在 AI 的迭代更新上，需要理解企业如何尽

量规避或削减 AI 可能的社会腐化效应，比如 AI 欺骗性对人的道德腐化以及 AI 有限性中人的干扰破坏倾向，并尽可能使 AI 技术以及 AI 应用朝着优化整体企业生态和提升社会福利的方向改进。AI 的级联效应是巨大的，这是因为 AI 在一定程度上模糊了组织边界，串联了更多主体，这可能有助于企业间的开放合作以及新型商业生态的培育（Bailey et al.，2022；Panait & Luke，2005）。此外，训练数据被广泛收集，算法模型渗透在各个行业，算法生成的结果（无论真实与否）又通过其真实的存在持续影响和塑造着社会现实（Kissinger et al.，2021；Lindebaum et al.，2023）。因此，未来研究可以关注由 AI 议题衔接起的企业、监管机构、非营利组织、社会活动家、技术领袖等多元角色之间，如何冲突对抗、协商斡旋或砥砺合作以共同定义合理化 AI 的内涵，构建负责任 AI 的制度，并采取对齐行动（Faulconbridge et al.，2023）。

综合来看，AI 基于大数据的判断能力可能辅助企业的战略决策，研究可以进一步讨论 AI 辅助对不同类型的决策如定位、竞争、创新、联盟，以及不同环境情境如高复杂度或不确定性程度下决策的影响。AI 在信息搜寻处理方面的综合性和及时性，可能对创业创新机会的发掘和利用有所帮助。未来研究可以考虑这种可被 AI 发现的商业机会的隐藏性、新颖性以及可操作性。最后，当 AI 作为新型智能员工嵌入现有的组织运营时，也将不可避免地影响组织内部的权力生态、分工方式、知识结构、制度文化及其流动方向。我们期待一个 AI 增强型组织的出现，但同时也需要对 AI 潜在的缺陷与误区予以足够的重视和应对。

接受编辑：主编团队

收稿日期：2024 年 11 月 7 日

接受日期：2024 年 11 月 22 日

作者简介

苏孟玥，西交利物浦大学和谐管理研究中心管理学助理教授。于 2023 年在北京大学国家发展研究院获得管理学博士学位。研究兴趣涉及组织学习、知识管理、动态能力和企业社会责任等领域。其研究成果发表于 *Organizational Dynamics* 和 *Academy of Management Best Paper Proceedings*。2023 年获得美国管理学会（AOM）Strategic Management（STR）分会的杰出论文奖，2024 年获得美国管理学会（AOM）Communication，Digital Technology，and Organization（CTO）分会的最佳论文奖。

马浩（通讯作者，E－mail：ma@ bimba. pku. edu. cn），北京大学国家发展研究院管理学教授，发树讲席教授，BiMBA 商学院学术主任。于 1994 年在得克萨斯大学奥斯汀校区商学院获得战略管理博士学位。研究兴趣涉及战略管理、创业创新和国际商务等领域。其研究成果发表于 *Academy of Management Review*、*Journal of International Business Studies*、*Journal of Business Venturing* 等学术期刊以及 *Organizational Dynamics* 和 *Business Horizons* 等面向管理实践者的期刊。曾出版中英文管理学著作 20 余本。为全球近百家各类企业提供管理咨询、定制研究与高管培训。

参考文献

［1］Anthony，C.，Bechky，B. A.，& Fayard，A. L.

2023. "Collaborating" with AI: Taking a system view to explore the future of work. *Organization Science*, 34: 1672-1694.

[2] Arias-Pérez, J., & Vélez-Jaramillo, J. 2022. Understanding knowledge hiding under technological turbulence caused by artificial intelligence and robotics. *Journal of Knowledge Management*, 26: 1476-1491.

[3] Bailey, D. E., Faraj, S., Hinds, P. J., Leonardi, P. M., & von Krogh, G. 2022. We are all theorists of technology now: A relational perspective on emerging technology and organizing. *Organization Science*, 33: 1-18.

[4] Balasubramanian, N., Ye, Y., & Xu, M. 2022. Substituting human decision-making with machine learning: Implications for organizational learning. *Academy of Management Review*, 47: 448-465.

[5] Bengio, Y., Hinton, G., Yao, A., Song, D., Abbeel, P., Darrell, T., Harari, Y. N., Zhang, Y. Q., Xue, L., Shalev-Shwartz, S. & Hadfield, G. 2024. Managing extreme AI risks amid rapid progress. *Science*, 384: 842-845.

[6] Bertini, M., & Koenigsberg, O. 2021. The pitfalls of pricing algorithms: Be mindful of how they can hurt your brand. *Harvard Business Review*, 99: 74-83.

[7] Chamorro-Premuzic, T., & Ahmetoglu, G. 2016. The pros and cons of robot managers. *Harvard Business Review*, 12: 2-5.

[8] Choudhary, V., Marchetti, A., Shrestha, Y. R., & Puranam, P. 2023. Human-AI Ensembles: When Can They Work? *Journal of Management*, https://doi.org/10.1177/01492063231194968.

[9] Collina, L., Sayyadi, M., & Provitera, M. 2023. Critical issues about AI accountability answered. *California Management Review Insights*.

[10] Connelly, C. E., Zweig, D., Webster, J., & Trougakos, J. P. 2012. Knowledge hiding in organizations. *Journal of Organizational Behavior*, 33: 64-88.

[11] De Cremer, D., & Kasparov, G. 2021. AI should augment human intelligence, not replace it. *Harvard Business Review*, 18: 1.

[12] Dupuy, J. 2024. *From Taylor Swift to troll farms: AI's real impact on election* 2024 *URL* https://www.forbes.com/sites/joshuadupuy/2024/10/31/from-taylor-swift-to-troll-farms-ais-real-impact-on-election-2024/.

[13] Ezzamel, M., Hyndman, N. S., Johnsen, Å., Lapsley, I., & Pallot, J. 2004. Has devolution increased democratic accountability? . *Public Money & Management*, 24: 145-152.

[14] Faulconbridge, J., Sarwar, A., & Spring, M. 2023. How professionals adapt to artificial intelligence: The role of intertwined boundary work. *Journal of Management Studies*, https://doi.org/10.1111/joms.12936.

[15] Gatrell, C., Muzio, D., Post, C., & Wickert, C. 2024. Here, there and everywhere: On the responsible use of artificial intelligence (AI) in management research and the peer-review process. *Journal of Management Studies*, 61: 739-751.

[16] Giroux, M., Kim, J., Lee, J. C. & Park, J. 2022. Artificial intelligence and declined guilt: Retailing morality comparison between human and AI. *Journal of Business Ethics*, 178: 1027-1041.

[17] Glikson, E., & Woolley, A. W. 2020. Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, 14: 627-660.

[18] Gregory, R. W., Henfridsson, O., Kaganer, E., & Kyriakou, H. 2021. The role of artificial intelligence and data network effects for creating user value. *Academy of Management Review*, 46: 534-551.

[19] Heaven, W. D. 2020. Google's medical AI was

super accurate in a lab. Real life was a different story. *MIT Technology Review*, 4: 27.

[20] Jia, N., Luo, X., Fang, Z., & Liao, C. 2024. When and how artificial intelligence augments employee creativity. *Academy of Management Journal*, 67: 5-32.

[21] Karunakaran, A., Orlikowski, W. J., & Scott, S. V. 2022. Crowd-based accountability: Examining how social media commentary reconfigures organizational accountability. *Organization Science*, 33: 170-193.

[22] Kemp, A. 2024. Competitive Advantage through Artificial Intelligence: Toward a Theory of Situated AI. *Academy of Management Review*, https://doi.org/10.5465/amr.2020.0205.

[23] Kissinger, H. A., Schmidt, E., & Huttenlocher, D. 2021. *The Age of AI: and Our Human Future*. Hachette UK.

[24] Köbis, N., Bonnefon, J. F., & Rahwan, I. 2021. Bad machines corrupt good morals. *Nature Human Behaviour*, 5: 679-685.

[25] Lindebaum, D., Moser, C., & Islam, G. 2023. Big Data, Proxies, Algorithmic Decision-Making and the Future of Management Theory. *Journal of Management Studies*, https://doi.org/10.1111/joms.13032,

[26] Ma, H., & Su, M. 2024a. Whom to Sue? Liability of Unaccountability in AI Organizational Management. *Working Paper*.

[27] Ma, H., & Su, M. 2024b. The bounded intelligence of AI: Superficiality and deceivability. *Organizational Dynamics*, 101100.

[28] Ma, H., & Su, M. 2024c. From mutual deceit to mutual enhancement: The positive reinforcement of human-AI relationship. *Working Paper*.

[29] Ma, H., & Su, M. 2024d. Artificial stupidity and coping strategies. *Organizational Dynamics*, 101059.

[30] Metcalf, L., Askay, D. A., & Rosenberg, L. B. 2019. Keeping humans in the loop: Pooling knowledge through artificial swarm intelligence to improve business decision making. *California Management Review*, 61: 84-109.

[31] Mirzadeh, I., Alizadeh, K., Shahrokhi, H., Tuzel, O., Bengio, S., & Farajtabar, M. 2024. Gsm-symbolic: Understanding the limitations of mathematical reasoning in large language models. *arXiv preprint arXiv*, 2410.05229.

[32] Murray, A., Rhymer, J. E. N., & Sirmon, D. G. 2021. Humans and technology: Forms of conjoined agency in organizations. *Academy of Management Review*, 46: 552-571.

[33] Ng, A. 2016. What artificial intelligence can and can't do right now. *Harvard Business Review*, 9: 1-4.

[34] Omidvar, O., Safavi, M., & Glaser, V. L. 2023. Algorithmic routines and dynamic inertia: How organizations avoid adapting to changes in the environment. *Journal of Management Studies*, 60: 313-345.

[35] Pakarinen, P., & Huising, R. 2023. Relational expertise: What machines can't know. *Journal of Management Studies*, https://doi.org/10.1111/joms.12915.

[36] Panait, L., & Luke, S. 2005. Cooperative multi-agent learning: The state of the art. *Autonomous Agents and Multi-agent Systems*, 11: 387-434.

[37] Park, P. S., Goldstein, S., O'Gara, A., Chen, M., & Hendrycks, D. 2024. AI deception: A survey of examples, risks, and potential solutions. *Patterns*, 5: 5.

[38] Raisch, S., & Fomina, K. 2023. Combining human and artificial intelligence: Hybrid problem-solving in organizations. *Academy of Management Review*, https://doi.org/10.5465/amr.2021.0421.

[39] Raisch, S., & Krakowski, S. 2021. Artificial

intelligence and management: The automation-augmentation paradox. *Academy of Management Review*, 46: 192-210.

[40] SEC. 2024. SEC *charges two investment advisers with making false and misleading statements about their use of artificial intelligence*. https://www.sec.gov/newsroom/press-releases/2024-36.

[41] Selbst, A. D., & Barocas, S. 2018. The intuitive appeal of explainable machines. *Fordham Law Review*, 87: 1085-1139.

[42] Spring, M. 2024. *How X users can earn thousands from US election misinformation and AI images*. https://www.bbc.com/news/articles/cx2dpj485nno.

[43] Staw, B. M., Sandelands, L. E., & Dutton, J. E. 1981. Threat rigidity effects in organizational behavior: A multilevel analysis. *Administrative Science Quarterly*, 501-524.

[44] Vanneste, B. S., & Puranam, P. 2024. Artificial Intelligence, Trust, and Perceptions of Agency. *Academy of Management Review*, https://doi.org/10.5465/amr.2022.0041.

[45] Yagoda, M. 2024. *Airline held liable for its chatbot giving passenger bad advice-what this means for travelers*. https://www.bbc.com/travel/article/20240222-air-canada-chatbot-misinformation-what-travellers-should-know.

[46] Yam, K. C., Tan, T., Jackson, J. C., Shariff, A., & Gray, K. 2023. Cultural Differences in People's Reactions and Applications of Robots, Algorithms, and Artificial Intelligence. *Management and Organization Review*, 19: 859-875.

[47] Zuboff, S. 2019. Surveillance capitalism and the challenge of collective action. *New Labor Forum*, 28: 10-29.

[48] Zuboff, S. 2022. Surveillance capitalism or democracy? The death match of institutional orders and the politics of knowledge in our information civilization. *Organization Theory*, 3: 1-79.

The Dark Side of AI in Organizational Management: An ABCD Framework

Mengyue Su[1] Hao Ma[2]

(1. Xi'an Jiaotong-Liverpool University, HeXie Management Research Centre;
2. Peking University, National School of Development)

Abstract: Despite the sustained attention and keen expectations surrounding the application of Artificial Intelligence (AI) in organizational management, it is imperative to recognize the dark side and potential negative implications associated with AI. Such a recognition will enable a more comprehensive, objective, and precise evaluation of AI's full potential within the sphere of organizational management. Accordingly, this paper advances an overarching analytical framework, referred to as ABCD, which hinges specifically on four major aspects of the potential drawbacks and problems of AI: Accountability, Boundedness, Cheating, and *Dumbness*. At the factual judgment level, boundedness and dumbness examine the systematic deficiencies in AI's capabilities. At the value judgment level, accountability and cheating scrutinize AI's flaws in legal and ethical dimensions. Specifically, first, AI lacks accountability and could not be held accountable. Naturally, AI's unexplanability makes it difficult for the assignment of responsibilities to specific decision-makers, humans or machines, in case of major decision failures. Moreover, even when AI is found to be responsible, it will not be able to bear the ultimate responsibility as it cannot be punished legally, financially, or mentally. Second, so long as AI develops its intelligence based on data drawn from human knowledge and expertise, it will suffer the problem of bounded intelligence, as determined by a host of natural and technological barriers that prevent AI from fully capturing the training data as well as artificial barriers created deliberately by the human actors to protect their own interests. Third, just as human actors could cheat and manipulate AI when providing training data or engaging in AI's design and application, AI has been found to cheat humans, for whatever reasons. Finally, in the extreme case, AI could turn squarely into artificial dumbness and might potentially cause fatal errors and unfortunate disasters, as it categorically and tyrannically overrides human intelligence and suppresses human initiatives. The paper concludes with an examination of ABCD's implications for future research and management practice. In sum, while we are sanguine about the supposedly bright future of applying AI in organizational management, we must also take into account its dark side for a more balanced and reasoned account.

Key Words: accountability; bounded intelligence; cheating; artificial dumbness; artificial intelligence