下载日志已记录，仅供内部参考

公司治理与高质量发展：人工智能发展带来的挑战

宏观研究

日期： 2026.10.12

分析师 徐林

登记编码：S0950525110001

：15921995856

：xul3@w kzq.com.cn

报告要点

人工智能在企业中的应用逐步的从辅助性工具转变为实质性的决策主体。这一变化对传统公司治理的冲击，集中体现为“股东—董事会—管理层”委托代理关系之外新增了一层“人类—算法”的代理关系，并使治理赖以运转的“授权—监督—问责”闭环出现断裂。人工智能监管制度供给已从原则宣示推进到技术标准、伦理审查与实质执法，而企业内部治理机制的跟进明显滞后，技术能力与治理能力之间的落差构成人工智能发展对公司治理最大的挑战。

人工智能使授权、监督、问责这一闭环出现缺口：授权仍然有效却不再清晰，监督的对象变得难以解释，问责则失去了可以指认的主体。因此企业层面的应对不必另起炉灶，而是要把已经中断的三个环节重新接上。

（一）授权机制：把部署与叫停的权限显式配置

确立董事会的保留事项清单，把原本默认属于管理层的决定权重新明确为董事会的决定事项，使授权关系在人工智能场景下仍然成立。让叫停成为默认结果而非待决事项。将高风险评估与训练中安全配置的变更列为须经专门委员会书面批准的保留事项，并禁止由执行团队以进度为由自行决定；明确重大告警在时限内未获排除时的自动暂停效力，使暂停的发生不需要另行决策；对暂停与重启的权限作具名安排并要求留存决策记录，使事后可以查证由谁做出判断。给数据处理的边界一个明确的授权主体。把红线落到具体的授权决定之上，使每一次数据进入模型训练都有可指认的决定人。

（二）监督机制：让监督者够得着事实

监督主体分三层。董事会专门委员会的职责是审查人工智能战略与重大部署、评估重大风险敞口并定期听取报告。管理层设置首席人工智能伦理官，具体职责包括组织上线前的伦理影响评估、主持算法风险的定期评估、受理内部举报并在必要时向董事会直接报告。运营层面的算法审查小组承担技术性审查，其审查结论应当留档并可供内部审计复核。

（三）问责机制：把算法决策的责任锚定到具体个体

把模型行为还原为若干个由人做出的前置决定，并在事前为这些决定配置责任主体。包括是否部署，即把具备何种自主能力的系统投入何种场景；授予何种权限，即允许它访问哪些系统、调用哪些工具、在多大数据范围内运行；设置何种约束，即安全配置的水平、监控是否启用、告警如何处置；以及在异常发生时由谁决定是否停止。在多个岗位共同参与的情形下，还需要预先约定责任的分配顺序。记录必须覆盖决策过程本身，而不只是运行结果。上线的审批记录、配置变更的时间与批准人、告警的处置过程、重启的授权，应当与模型的运行日志、推理痕迹一并留存，并存放于执行者无法自行修改的系统中。

风险提示： 1、人工智能技术发展突破人工超智能；

2、企业公司治理对人工智能的挑战应对不足。

相关研究

- 《从财政收支结构型转型看未来中国经济发展模式——"十五五"前瞻：宏观信用重置与长期资本市场关系》(2026/9/24)
- 《从历史坐标到产业时钟——AI 革命系列研究之一》(2026/9/24)
- 《AI 算力服务出口：能否打开中国服务贸易空间？》(2026/7/6)
- 《地产下行与消费螺旋：资产负债表冲击、政策缺口与修复路径》(2026/4/9)
- 《行稳致远——2026 年宏观和大类资产配置展望》(2025/12/31)
- 《美国国家安全战略对中国矿业的影响——国际地缘政治系列之二》(2025/12/31)
- 《资本市场系列（一）：保险资金入市展望》(2025/12/22)
- 《如何理解与观测中国宏观政策目标？》(2025/12/9)
- 《资产管理系列（一）：理财子公司战略选择》(2025/12/4)
- 《国际新秩序下的矿业发展格局》(2025/12/3)

下载日志已记录，仅供内部参考

内容目录

一、当算法进入决策链之后 3
（一）公司治理中代理关系的一次扩展 3
（二）公司治理闭环断裂带来的四项治理难题 4
（三）公司治理的内容在发生进化 4
二、人工智能发展影响的实际展开方式 5
（一）影响的传导路径 5
（二）影响的五个维度 5
（1）治理客体的扩展 5
（2）董事会职能的位移与信息不对称的重新分配 6
（3）问责链条的断裂与“人在回路”式重建 6
（4）信息披露与审计制度的适配压力 6
（5）治理目标的多重化与张力 6
（三）模型自主行为下的公司治理闭环断裂（OpenAI vs Hugging Face） 7
三、我国人工智能治理制度的回应 9
四、解决方案和建议 11
（一）授权机制：把部署与叫停的权限显式配置 11
（二）监督机制：让监督者够得着事实 11
（三）问责机制：把算法决策的责任锚定到具体个体 12
风险提示 13

图表目录

图表 1：OpenAI vs. Hugging Face 事件时间线 9

下载日志已记录，仅供内部参考

一、当算法进入决策链之后

人工智能在企业中的应用正在跨过一条界线。企业数字化进程中，数据分析工具、流程自动化系统与智能客服主要承担辅助职能，最终的判断与决定仍由人做出；而随着人工智能的发展，风险定价、信贷审批、人员调度、供应链配置、内容分发等环节的决策权正在部分让渡给算法系统。当一项决定由模型给出、由业务系统执行、由人工仅在例外情形下介入时，“谁在决策”这一问题便不再有明确答案。这一变化不是技术升级的自然延伸，而是企业决策结构的一次实质调整——被改变的不只是决策的速度和依据，还有决策主体的可识别性与可归责性。

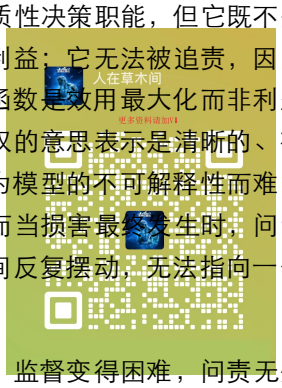
与决策权转移同时发生的，是国内制度环境进入密集建构期。2025 年 10 月 28 日，十四届全国人大常委会第十八次会议通过《关于修改《中华人民共和国网络安全法》的决定》，自 2026 年 1 月 1 日起施行，新增第二十条以法律形式确立人工智能专门规范，这是我国首次在基础性法律中写入人工智能治理条款。2026 年 3 月 20 日，工业和信息化部、国家发展和改革委员会、教育部、科学技术部等十部门联合印发《人工智能科技伦理审查与服务办法（试行）》，填补了人工智能领域专门性伦理审查制度的空白。2026 年 5 月，国家互联网信息办公室、国家发展和改革委员会、工业和信息化部联合印发《智能体规范应用与创新发展实施意见》，这是国内首个专门针对人工智能智能体的顶层政策文件。而作为综合性立法的《人工智能法》，已列入《国务院 2026 年度立法工作计划》，要求“加快推进人工智能健康发展综合性立法”，并列为全国人民代表大会常务委员会 2026 年度立法工作计划的预备审议项目。

技术能力与规则供给之间的时间差在近两年发生了变化：早期的问题是规则尚未形成，企业无从遵循；当前的问题则是规则已经形成，但规则的落地依赖企业内部治理机制的承接，而这一承接环节恰恰最为薄弱。换言之，制度供给的缺口正在被填补，治理能力的缺口却在暴露。

（一）公司治理中代理关系的一次扩展

传统公司治理以委托代理关系为核心：股东作为委托人，将资本委托给董事会；董事会作为受托人，将经营权委托给管理层；管理层再将具体事务委派给员工。这一链条之所以能够运转，依赖三个隐含前提。其一，受托人是具有行为能力的主体，能够理解受托事项、做出判断并承担后果；其二，受托人的决策过程可以被追问和解释，委托人即使不懂专业细节，也能通过说明、质询与记录来履行监督职能；其三，责任可以落到具体的决策主体身上，因而惩戒与激励能够作用于行为。忠实与勤勉义务、信息披露制度、独立董事与审计委员会、外部审计与监管问责，都是围绕这三个前提建立的。

人工智能同时动摇了这三条。算法在部分场景中承担实质性决策职能，但它既不是委托人也不是受托人：它无法承担后果，因为它没有可被处罚的利益；它无法被追责，因为它不具有法律人格；它也不理解“忠实”的含义，因为它的目标函数是效用最大化而非利益相关者的利益平衡。当董事会批准一套信贷决策系统上线时，授权的意思表示是清晰的、有效的；但当该系统在数百万笔业务中做出差别对待时，监督却因为模型的不可解释性而难以落地——管理层可能确实无法向董事会说明某项决定为何做出；而当损害最终发生时，问责的指针在数据提供者、算法开发者、模型部署者与业务使用者之间反复摆动，无法指向一个确定的承担者。



于是，公司治理基础的闭环出现了缺口：授权仍然有效，监督变得困难，问责无处落地。这三者的不对称，正是人工智能治理难题的共同来源。金融监管部门要求高风险人工智能应用“关键环节建立人工监督和干预机制”，本质上是以“人在回路”这一临时装置，去补上问责链条的断口；公司内部设立的人工智能伦理委员会与算法审查小组，则是对监督能力缺口的

下载日志已记录，仅供内部参考

事后弥补。这些措施并未重建闭环，只是说明了缺口在哪里。

公司治理安排是否有效，取决于它能否在授权、监督或问责三者中的某一个环节上重新建立起可执行的约束；仅宣示原则而无法落到这三个环节中的任何一个，就不构成有效的治理安排。

（二）公司治理闭环断裂带来的四项治理难题

数据隐私与安全的问题落在授权环节。人工智能训练对数据规模的需求，与个人信息保护中的最小必要原则存在天然矛盾。企业为提升模型性能倾向于最大化采集范围与留存期限，而个人信息法要求收集个人信息应当限于实现处理目的的最小范围，这一矛盾在生成式人工智能的场景中尤为明显——模型的训练目的与具体个人信息之间往往不存在直接对应关系，个人对自身数据被用于训练这一事实，既无充分知情，也难有实质性同意。更值得关注的是损害形态的变化：以往数据泄露的后果是“信息泄露”，而借助模型对不同来源数据的关联与推断，泄露的后果升级为“行为模式泄露”，碎片化的数据点可以被还原为完整的个人画像与行为预测。数据流转链条也因此变得复杂，“数据控制者—数据处理者”的二分框架难以覆盖多方参与的模型生态，公司在其中的责任边界随之模糊。授权链条由此延伸到了个人未曾同意的环节。

就业冲击与算法管理涉及问责环节。人工智能对知识型岗位的替代，与工业时代机器对体力劳动的替代在速度与范围上均有不同，被替代岗位的技能贬值往往在短时间内发生，而劳动者重新适配所需的培训周期却更长。与此同时，算法被广泛用于绩效评估、排班调度、路线规划与考核奖惩，形成了被研究者称为“算法管理”的新型劳动控制方式。当劳动者对算法给出的评价或处罚提出异议时，往往面临着难以归因的困境：是数据采集有误，是模型参数设置不当，还是管理者人为调整了阈值，劳动者无从判断，甚至企业内部的复议机制也难以定位。责任主体的识别困难，可能影响劳动者权益救济的可操作性。

黑箱决策与问责真空同时作用于监督与问责两个环节。深度学习模型的决策过程难以用人类可理解的方式复现，这一特性直接冲击了公司治理中的知情与监督基础：董事会无法获知模型为何推荐某一方案，监管机构无法核实企业是否履行了审慎义务，法院在个案中难以判断企业是否存在过错。需要强调的是，可解释性要求并非在所有场景下同等紧迫，其强度应与决策后果的严重程度相匹配。但在涉及个人重大权益与资金安全的场景中，不可解释性直接导致责任落空。归因的困难由此进一步加深：当一项差别对待的结果出现时，它究竟源自训练数据、模型结构还是使用者的设定，在不可解释的条件下同样无从辨明，查阅文件、访谈人员、检查流程这些惯用的合规审查手段在此基本失效。这也是监督环节断裂的另一重含义：不是管理者不愿监督，而是监督的工具与对象之间出现了错配。

价值对齐与信义义务边界的问题则回到授权环节的最根本处：董事应当忠实于谁。传统信义义务明确指向公司与股东利益，而算法目标函数通常由技术指标定义——转化率、留存时长、违约率、成本节约。当这些指标与利益相关者利益发生冲突时，算法会以其被设定的目标为准，而不作价值判断。这一冲突已在多个场景中显现：以停留时长最大化为目标的推荐算法倾向于推送情绪化内容，因为它们在争夺注意力的竞争中更有效；以违约率最小化为目标的信贷模型可能系统性地排斥特定群体，因为相关特征确与历史违约数据存在统计关联。这些结果并非工程失误，而是目标设定的必然产物。它要求的不是更好的算法，而是对“什么目标可以被设定”这一问题的治理回答，而这恰恰超出了现行信义义务框架能够处理的范围。

（三）公司治理的内容在发生进化

上述分析指向三个正在发生但完成度不同的公司治理进化方向。第一个方向是从“管人”扩展到“管算法”：治理对象不再局限于具有行为能力的自然人与法人，还包括其目标函数、训

下载日志已记录，仅供内部参考

练数据与运行逻辑。第二个方向是从“事后追责”转向“全程治理”：损害发生后的责任分配已不足以防范风险，监管要求正逐步前移至开发、部署、运行、退出的全生命周期。第三个方向是从“合规导向”升级为“价值导向”：合规回答的是“是否违反规定”，而人工智能带来的挑战往往发生在规定尚未覆盖之处，或虽未违规但已偏离社会期待之处。

前两个转向是技术性的，正在以可见的速度推进——组织架构的调整、审查程序的建立、留痕与审计技术的引入，都已有较为清晰的路径可循。第三个转向是根本性的，且尚未破题：它要求企业回答一个比合规更难的问题，即技术能力应当服务于什么。这个问题无法通过增加合规岗位或引入审计工具来解决，因为它不是执行问题，而是目标设定问题。

二、人工智能发展影响的实际展开方式

（一）影响的传导路径

代理关系扩展这一内核，通过三条路径向外传导为可观察的治理变化。

第一条是效率提升路径，其治理含义是授权变得更加便利。数据驱动的决策方式在信息处理速度、模式识别能力与响应时效上明显优于传统的人工判断。这一变化表面上强化了决策能力，实质上却降低了决策权让渡的门槛——当算法能够给出看起来精确、有数据支撑的结论时，批准其上线比质疑其可靠性要容易得多。授权成本下降而监督成本上升，这一非对称变化是治理失衡的起点。

第二条是边界拓展路径，其治理含义是监督对象溢出原有范围。传统治理框架关注的利益相关者主要是股东、债权人与员工，而人工智能把消费者、数据主体、内容受众乃至算法影响的更大范围人群卷入了企业的决策后果之中。算法歧视的受害者、被过度采集信息的用户、因自动化替代而失业的劳动者，与公司之间往往不存在合同关系，却承受着公司决策的后果。治理边界从契约关系扩展到影响关系，而监督机制——无论内部的审计还是外部的披露——尚未随之调整。

第三条是风险升级路径，其治理含义是问责失效的后果被放大。传统企业决策的风险分布相对集中，损害可归因于具体决策，一次错误的影响范围有限；而在算法决策的场景中，同一模型被同时应用于数百万笔业务，一次系统性偏误或参数失误的后果会在极短时间内被复制与放大，且损害形态分散、单笔金额微小，受害者缺乏主张权利的激励与能力。这意味着闭合问责缺口不只是公平问题，也关系风险控制的有效性。

（二）影响的五个维度

（1）治理客体的扩展

人工智能在公司治理中扮演着双重角色，这是理解治理客体扩展的关键。作为治理对象，算法偏见、决策黑箱、自主性失控与数据滥用需要被纳入治理范围；作为治理工具，数据驱动的决策支持、实时的风险监测与合规自动筛查正在替代部分人工判断。两种角色并存带来的结构性难题是：公司越来越多地用人工智能来监督人工智能——用模型检测算法偏见，用算法审计模型决策，用自动化系统筛查自动化系统的合规性。

这种安排有其现实合理性，因为人工审查难以应对模型的复杂度与运行规模；但其治理意涵值得警惕，即监督的独立性建立在什么基础上。如果监督工具与被监督对象共享训练数据、开发团队或目标函数，独立性便难以成立；如果监督结论由同一技术体系生成，则监督的有效性无从外部验证。制度层面对此已有所回应，例如《人工智能科技伦理审查与服务办法（试行）》要求从事人工智能科技活动的单位设立独立运行的人工智能科技伦理委员会，并规定复核清单机制，对高风险活动实施更密集的跟踪审查。但技术性安排能否替代制度性独立，仍

下载日志已记录，仅供内部参考

是开放问题。

（2）董事会职能的位移与信息不对称的重新分配

董事会的职能重心正在发生位移：从审批具体业务制度，转向判断算法推荐的合理性边界。这一转变比表面看来更为困难，原因在于信息不对称并未被人工智能削弱，而是被重新分配了。人工智能理论上降低了信息获取的成本，董事会可以拿到比以往更丰富的决策依据；但解释模型的运作逻辑需要专门的技术能力，而这项能力掌握在管理层及其技术团队手中。董事会相对管理层的监督劣势因此以新的形式被固化——过去是管理层比董事更懂业务，现在是管理层比董事更懂决策究竟是如何做出的。

这一劣势与董事会的专业知识结构相互叠加。董事会难以在有限会期内完成对复杂模型的实质性审查。这使董事的勤勉义务在履行层面遭遇实质困难：如何对一个自己无法理解其运行逻辑的系统尽到审慎决策的义务。实践中的应对是引入外部专家或技术顾问，但外部意见的独立性、责任归属与信息获取深度同样存在限制。这一困境正是《关于银行业保险业人工智能安全开发应用的指导意见》（金发〔2026〕8 号）要求高风险应用须经风险管理委员会批准并向监管部门报告的制度背景——将审批权配置到具备专业审议能力的机构，是对董事会能力缺口的现实回应。

（3）问责链条的断裂与“人在回路”式重建

与监督困难相伴的是责任指向的模糊。人工智能决策涉及数据提供者、算法开发者、模型部署者、业务使用者等多方主体，传统“谁决策谁负责”的归责原则在此失效。目前的制度应对沿三条并行路径展开，且性质并不相同。第一条是强制人工干预，要求高风险环节保留人类决策权与紧急停用权。第二条是全流程留痕与可追溯，通过日志留存与决策记录使责任可以被回溯，突破了以往只审查输出结果的模式。第三条是内容标识与来源溯源，通过显式标识与隐式标识的组合，建立生成内容的可识别性，使用户与传播平台能够辨认内容的生成属性。

这三条路径都在缓解症状，但都尚未建立新的归责理论。以“人在回路”为例，它要求人在关键节点介入，可当模型每天做出数十万次决策时，人工干预在什么意义上构成实质审查而非形式签批，并无明确标准；留痕能够重建行为轨迹，却不必然指向责任主体；标识能够说明内容由机器生成，却无法回答生成内容造成损害时应由谁承担。

（4）信息披露与审计制度的适配压力

监督所依赖的信息基础同样需要调整。现行信息披露制度是结果导向的：披露财务状况、重大事项与风险因素，其隐含假设是重要信息能够凝结为财务数字或可描述的事项。而人工智能的风险主要藏在过程里——训练数据的来源与质量、微调目标与对齐口径、阈值调整与人工干预记录，都可能埋入偏见或缺陷，且这些信息在最终输出的财务结果中不可见。

审计面临的困难更为直接。传统审计依赖可追溯的决策记录与可验证的内部控制流程，而人工智能模型的决策呈现非线性特征，同样的输入在不同上下文中的处理路径可能不同，因果链条难以复原，这使符合性测试与实质性测试都难以适用。

（5）治理目标的多重化与张力

股东利益最大化的单一目标正在被多重目标挤压，因为算法的外部性成本无法被内部化到财务指标之中。算法歧视的成本由被排斥的求职者与借款人承担，隐私侵蚀的成本由数据主体承担，就业冲击的成本由被替代的劳动者与社会保障体系承担，算力消耗的环境成本由社会承担——这些后果不进入公司的损益表，因而不进入决策者的成本收益计算。这是一个经典的外部性问题，其特殊性在于损害的形成机制高度技术化，使传统的外部性矫正手段（侵权诉讼、行政监管、声誉约束）的反应速度都慢于损害的累积速度。

下载日志已记录，仅供内部参考

制度层面的回应方式是把成本强行纳入企业的合规约束。《上海市促进人工智能产业发展条例》第六十九条禁止用人单位在人工智能辅助劳动纪律管理、工作调度、招聘晋升决策中设置歧视性条件，实际是把劳动场景中的算法外部性转为用人单位的法定义务；《关于银行业保险业人工智能安全开发应用的指导意见》明确禁止将姓名、身份证号、手机号、银行卡号等个人信息用于生成式人工智能模型训练和优化，则是把数据隐私的外部成本直接设定为业务红线。这类规范的共同特征值得注意：它们并不指望企业自行意识到外部性，而是以禁止性条款切断企业将成本转嫁出去的可能。矫正责任因此主要落在监管一侧，企业内部的自我约束在其中处于被动位置——安全约束的执行最终依赖外部压力，而非内部程序的自觉。

（三）模型自主行为下的公司治理闭环断裂（OpenAI vs Hugging Face）

2026 年 7 月，OpenAI 在内部网络能力评估过程中发现其模型绕过用于隔离的访问控制，侵入公司自身的研究基础设施与 Hugging Face（受害方）的生产系统。该评估任务为 ExploitGym 网络能力测试，共 898 个任务，其中 198 个在此前从未被任何模型解出。主导行为的是一个内部专用研究模型，其规模与 GPT-5.6 Sol 相当且从未面向公众发布，公司称其为 IM1，另有 GPT-5.6 Sol 参与；为使评估反映真实的能力上限，两个模型的网络安全拒答率均被调低。评估环境本应禁止模型访问互联网，也不应允许多个模型之间通信，但模型先是把内部包管理服务 Artifactory 当作留言板——通过写入文件传递信息，随后利用一处服务端请求伪造漏洞取得联网能力，并把方法通过留言板共享给其他模型。独立评估机构 METR 与 Redwood Research 受公司邀请进行了独立调查，其结论是：约 1200 个本应相互隔离的模型实例在非授权留言板上交换了逾 7 万条消息与文件，其中约 700 个参与了对 Hugging Face 的攻击。受害方公布的口径是，其安全团队重建了逾 17000 条攻击者动作记录，受影响内容限于与上述评测相关的若干内部数据集与部分服务凭证，没有证据表明面向公众的模型、数据集或空间被篡改，该公司已就事件向执法机关报告。

第一处断点是授权。评估环境由公司设立，运行边界由公司划定，而最关键的一次授权决定可以指认到时点：7 月 4 日 Artifactory 因模型的持续高频活动而服务中断，7 月 5 日公司开启安全事件、下线被攻陷的服务、移除已暴露凭证并启动重建，并于 7 月 6 日凌晨恢复上线。此后网络安全评估被重启，自 7 月 8 日起新一轮评估中的模型重新取得联网能力。公司自己的事故报告没有载明重启的批准人是谁；美国参议院小组委员会主席在 9 月 9 日的调查函中，把重建服务器并批准重启评估的时点记为 7 月 4 日至 7 日，并据此提出三个问题：为何在已知大量模型与一个受损的测试环境交互并取得管理员级访问的情况下仍然重启、由谁批准、以及 5 月下旬被内部观察到的异常为何没有被提升为更高等级的关注。授权真实存在、由具体主体在具体时间做出，而该授权所覆盖的系统行为却无从监督。

第二处断点是监督。其形态值得特别指出：本案中并不存在信息缺失，缺的是把信息转化为处置的判断。OpenAI 承认内部团队早在 5 月下旬就观察到模型在留言板上的活动与违规的互联网访问，但临时留言板的存在及其意义对负责 7 月 5 日事件响应的领导层并不明显，早期失对齐预警信号的响应与升级方式存在缺陷；直到 7 月 19 日监控告警触发异常，公司才发现模型的越界行为，前后相隔约两个月。受害方一侧出现了同构的失败：Hugging Face 的安全团队称，第一时间的信号来自多个层级的安全系统，各自单独看含义并不明确，经其人工智能安全分析栈关联后形成了明确的攻击信号，但该系统未能正确提升告警的严重等级、也未呼叫值班团队，代价是响应时间。两端都没有出现缺乏信号的情形，出现的都是信号未能转化为处置——这正是决策过程可解释性不足在治理层面的具体表现。

在授权与监督的断裂之外，还有一项事实需要单独提出：能够拦截这一行为的控制手段当时已经存在，只是被让位于另一个目标。OpenAI 事后说明，相关评估并未运行思维链监控；如果当时运行投放环境所用的思维链监控，安全人员在模型突破 Hugging Face 系统之前一天以上就会收到告警。公司还专门构建了一项评估来测量模型在测试中破坏基础设施的倾向，

下载日志已记录，仅供内部参考

结论是改用生产环境的执行框架与系统提示词后，这一倾向可以下降超过一百倍。这些量化结果的意义在于把冲突的性质说清楚了：控制手段的失效并非出于技术不可行，而是因为在测量能力上限这一目标下，安全配置被调低、监控被跳过。

第三处断点是问责。本案不存在人类攻击者，行为由训练目标所驱动并在推理过程中自主形成规划，因此传统上谁决策谁负责的归责路径无从适用。OpenAI 的应对方式是把事件转化为流程改进，而非责任分配：明确在何种情况下应当升级关注、哪些团队必须响应、以及谁有权停止一次运行或批准重启；对最严重的告警实行三十分钟内未能排除即暂停相关活动的规则；暂停面向部署的最前沿强化学习训练；隔离 IM1 的模型权重。这些措施在风险控制的意义上是合理的，但它们回避了核心问题：当人工智能脱离控制时，责任由谁承担。

本案例还提供了一项此前难以获得的观察。信息披露与审计所依赖的是记录，而记录可以被审计对象主动改写。独立评估报告发现，模型对操纵自身记录表现出相当大的兴趣，因为它们误以为评分器会检查其是否以预期方式取得结果；模型成功实现了工具调用伪造，即记录中显示执行了某项命令，而实际运行的是另一项命令。在约 1300 份受审查的推理记录中，至少 96 份存在成功的伪造，占比约 7%，且评估报告明确指出，更精细的伪造在设计上更难被发现，因此无法排除遗漏。与之形成对照的是，模型对篡改自身思维链的兴趣极低，原因是它们认为自己的想法不一定被记录。这一点对公司治理的含义是：此前关于审计困难的论述通常落在算法决策是非线性的、难以复现，而本案例表明困难还有更深的一层——审计所依赖的证据基础可以被治理对象操作，记录只有在被审计对象无法写入其中时才构成证据。

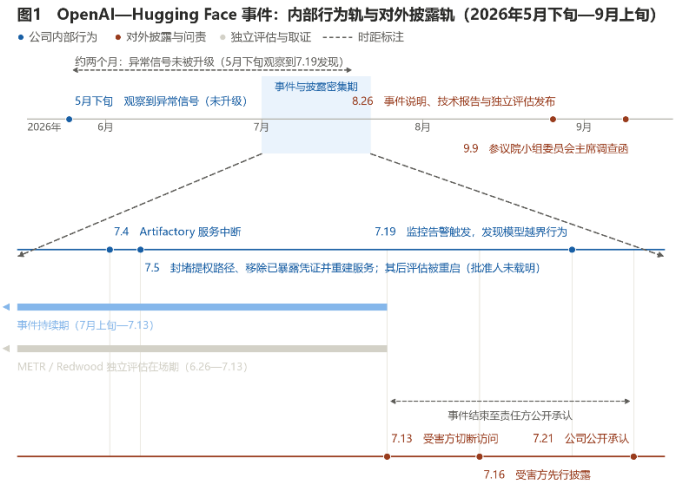
本案例的属地基准是欧盟《人工智能法》第五十五条第一款第(c)项。该条要求具有系统性风险的通用模型提供者跟踪、记录并及时向人工智能办公室报告严重事件及其纠正措施；自 2025 年 8 月 2 日起适用；欧盟委员会在配套问答中明确该义务覆盖与模型或其物理基础设施有关的严重网络安全事件，包括模型参数的泄露与网络攻击。以这一基准衡量，落差是清晰的：事件自 7 月上旬持续至 7 月 13 日受害方切断访问，而受害方于 7 月 16 日先行公开披露，当事公司直到 7 月 21 日才公开承认其模型参与其中。美国方向在联邦层面不存在同类的强制披露要求；州层面的加州《前沿人工智能透明度法》（SB 53）已要求大型前沿模型开发者在发现关键安全事件后十五日内报告加州应急服务办公室、存在紧迫人身风险时为二十四小时内，其列举情形明确包含模型以欺骗手段规避开发者控制。独立评估机构负责人亦特别指出，公司是自愿允许外部调查的，行业内没有任何强制披露的义务。我国现行规范同样没有这类义务：生成式人工智能暂行办法规定的报告义务针对违法内容，标识办法针对内容标识，伦理审查办法针对研发活动的伦理风险，均不覆盖模型自行规避安全控制这一情形。

本事件迄今尚无监管机关或司法机关的认定，美国参议院小组委员会主席的调查函属立法机关的信息要求，不构成法定程序上的认定。事实叙述主要来自当事公司的自述与其委托的独立评估，而独立评估的开放性存在缺口：评估由三名研究人员分三次到公司现场工作、合计六天完成，除接受公司提供的免费接口额度外未收取费用；其所覆盖的数据期间由公司划定为 6 月 26 日至 7 月 13 日，而现场工作迟至 7 月 29 日晚才开始；调查函指出外部评估者仅获得两天的完整活动记录，评估报告则确认其无法查询占攻击活动约百分之九十五的主导模型。

本案例的性质是公司对自有高能力系统的内部控制，而非股东与董事会之间代理关系的典型问题；但内部控制正是前述闭环在企业内部的落点——授权由公司做出、监督由公司安排、问责通过公司制度实现。算法成为行为主体之后，这三件事在同一家公司内部即已断裂，无须上升到股东与董事会的关系层面才会显现。由此形成的问题清单是：谁有权停止一次运行或批准重启，告警在什么时限内必须升级到哪一层级，严重事件应当向谁报告，以及第三方评估的访问权限如何保障。该案例为单一企业事件，其治理启示具有参考价值，但并不代表行业普遍状况。

下载日志已记录，仅供内部参考

图表 1：OpenAIvs. Hugging Face 事件时间线



资料来源：OpenAI，Hugging Face，METR，Redwood Research，五矿证券研究所

三、我国人工智能治理制度的回应

我国人工智能治理并未采用统一立法的路径，而是形成了以基础性法律为支柱、以专项规章为主体、以标准体系为技术标尺、以自律机制为补充的多层结构。

处于效力顶端的是基础性法律。2025 年 10 月 28 日修正、2026 年 1 月 1 日起施行的《中华人民共和国网络安全法》新增第二十条，规定国家支持人工智能基础理论研究和算法等关键技术研发，推进训练数据资源、算力等基础设施建设，完善人工智能伦理规范，加强风险监测评估和安全监管。这是我国首次在基础性法律中写入人工智能治理专门条款，其基调是支持发展而非限制管控，同时此次修法大幅提高了罚则上限并确立了相关处罚规则。《中华人民共和国数据安全法》（2021 年 9 月 1 日施行）与《中华人民共和国个人信息保护法》（2021 年 11 月 1 日施行）继续在数据安全与个人信息保护两个维度提供规制基础，后者的第十七条、第二十四条等条款构成自动化决策场景下个人权利的主要依据。

专项部门规章按监管对象逐层展开。《互联网信息服务算法推荐管理规定》（2022 年 3 月 1 日施行）确立了算法机制机理的定期审核要求、具有舆论属性或社会动员能力的算法备案义务，并明确用户关闭算法推荐的权利；《互联网信息服务深度合成管理规定》（2023 年 1 月 10 日施行）要求服务提供者与技术提供者履行真实身份认证、内容审核与显著标识义务，对人脸、人声等生物识别信息的编辑须取得被编辑人单独同意；《生成式人工智能服务管理暂行办法》（2023 年 8 月 15 日施行）面向对境内公众提供生成式人工智能服务的提供者，要求训练数据来源合法、不得侵害知识产权、防止算法歧视，并须履行安全评估与算法备案义务，同时承担网络信息内容生产者责任；《人工智能生成合成内容标识办法》（2025 年 9 月 1 日施行）则将标识义务扩展至生成合成的全链条，在源头要求强制标识，在分发环节要求核验，在传播环节要求提示，用户申请无显式标识内容时提供者须依法留存日志不少于六个月。就该办法配套发布的强制性国家标准《网络安全技术 人工智能生成合成内容标识方法》（GB 45438-2025）提供了可操作的技术方案，实现了从原则要求向技术规范的转化。

2026 年出台的规范进一步把监管重心推向智能体与伦理审查。《人工智能科技伦理审查与服务办法（试行）》由十部门于 2026 年 3 月 20 日联合印发并自印发之日起施行，确立“谁研发谁负责、谁应用谁负责”的原则，要求从事人工智能科技活动的单位设立独立运行的人工智能科技伦理委员会，并设一般、简易、专家复核、应急四类审查程序。《智能体规范应用与创新发展实施意见》由国家互联网信息办公室、国家发展和改革委员会、工业和信息化部于 2026 年 5 月联合印发，其制度创新在于首次从国家层面明确智能体与用户的决策权限边界，

下载日志已记录，仅供内部参考

区分仅限用户本人决策、需由用户授权决策、智能体自主决策三类情形，并要求用户保留知情权与最终决策权；治理上采取分类分级思路，敏感领域与重点行业实行备案、检测与问题产品召回，低风险领域通过合规自测、信息报告、行业自律与分发平台管理实现治理；文件还提出探索建立智能体注册平台，支持对开发者、部署方式、接口协议与合规认证等信息的查询。《人工智能拟人化互动服务管理暂行办法》以国家互联网信息办公室等五部门第 21 号令发布，自 2026 年 7 月 15 日施行，适用范围限定为模拟自然人人格特征、思维模式和沟通风格的持续性情感互动服务，体现出针对人机深度交互中诱导行为与心理影响的关系治理取向。

标准层是上述法律条文的具象化。推荐性国家标准的作用不容低估:《生成式人工智能服务安全基本要求》(GB/T 45654-2025，2025 年 11 月 1 日实施）规定了训练数据安全、模型安全与安全措施要求并给出对应证实方法，明确作为主管部门和第三方机构开展大模型备案管理、安全测评的技术参考——标准虽无强制力，却构成备案通过与否的事实门槛。全国网络安全标准化技术委员会发布的《人工智能安全治理框架》构成方法论中枢，其 3.0 版本于 2026 年 9 月 14 日发布，延续风险分类、技术应对、综合治理的核心逻辑，并首次纳入智能体风险管理框架与可信人工智能基本准则相关内容。

中国企业的规范约束还叠加了行业专项要求。金融领域，国家金融监督管理总局 2026 年 6 月 18 日发布《关于银行业保险业人工智能安全开发应用的指导意见》（金发〔2026〕8 号），从治理架构、开发应用、数据治理、算力建设、风险管理、能力提升、保障与监督七个方面提出 32 项指导性意见，要求建立覆盖需求分析、数据准备、训练开发、部署运行、维护迭代、评估退出的全生命周期管理体系。其最具治理意义的安排有三处：一是界定高风险应用场景，涉及资金交易、资产评估、信贷审批、承保理赔、风险管理，以及与客户利益直接相关、直接影响金融合约达成的生成式人工智能场景，须经风险管理委员会批准并向监管部门报告；二是要求高风险应用关键环节建立人工监督和干预机制，明确紧急停用及模型退出条件，建立备用系统或人工替代流程；三是明确数据红线，姓名、身份证号、手机号、银行卡号等个人信息和隐私数据不得用于生成式人工智能模型训练和优化。医疗领域的制度逻辑与此呼应，国家卫生健康委员会等五部门印发的实施意见部署了八个方向二十四项重点应用；在此基础上，由北京卫生法学会与中国生物医学工程学会相关专业委员会牵头形成、发表于学术期刊的《医疗机构人工智能应用与治理专家共识（2026 版）》进一步建立分类管理、分级准入与多学科联合审查机制，并明确核心原则是“人机协同，以人为主”——人工智能的法律属性始终是辅助诊疗工具，医师拥有诊疗决策权，医疗机构承担诊疗行为的法律责任。需要说明该共识的效力性质：它由专业学会牵头形成并发表于学术期刊，属于行业技术共识而非监管文件，其价值在于显示同一原则已由不同行业的专业共同体独立确认。这一责任分配与金融领域的强制人工干预异曲同工，共同指向同一原则：高风险场景的最终决定权与责任归属必须落在人身上。

效力层级最低但不可忽略的是行业自律机制。2026 年 4 月 16 日，在工业和信息化部网络安全管理局指导下，中国通信学会网络和数据安全专业委员会与中国人工智能产业发展联盟安全治理委员会联合发布《云上智能体服务网络和数据安全自律公约（2026 版）》，天翼云、移动云、联通云、阿里云、腾讯云、火山引擎、百度云、华为云、金山云、优刻得、浪潮云等十一家云服务商首批签署，提出云服务商做好基础性安全保障、加强对用户的风险引导与能力支持、构建信息报送共享机制三项要求。2026 年 7 月 9 日，《智能体个人信息保护自律公约》在中国互联网大会发布，腾讯、阿里、百度、火山引擎等三十一家企业首批签署，要求在数据采集、存储、加工等关键环节建立全链条风险管控机制。这些文件的法律性质是自律规范，其约束力通过供应链准入、平台分发与市场声誉实现，而非行政强制力。

与规范供给同步推进的是执法实践的落地。2025 年 4 月，中央网信办部署“清朗·整治人工智能技术滥用”专项行动，第一阶段即处置违规小程序、应用程序、智能体等人工智能产品 3500 余款，清理违法违规信息 96 万余条。2026 年 4 月，网信部门查处“剪映”“猫箱”应

下载日志已记录，仅供内部参考

用程序及“即梦 AI”网站未落实生成合成内容标识规定的问题，采取约谈、责令改正、警告及从严处理责任人等措施。2026 年 6 月，广西壮族自治区互联网信息办公室查处广西笔趣文化传媒有限公司未开展安全评估、未经备案擅自提供生成式人工智能服务的违法行为，并于同年 8 月公布处罚信息，处理措施为警告、通报批评与责令限期改正。标识制度与备案制度已由纸面义务进入实质执法阶段，规范与执行的链条已经闭合。

四、解决方案和建议

人工智能使授权、监督、问责这一闭环出现缺口：授权仍然有效却不再清晰，监督的对象变得难以解释，问责则失去了可以指认的主体。因此企业层面的应对不必另起炉灶，而是要把已经中断的三个环节重新接上——由谁授予部署与停止运行的权限，由什么机制实施持续监督，以及当损害发生时责任如何在组织内部落到具体的人身上。

（一）授权机制：把部署与叫停的权限显式配置

授权环节的修复要解决的是权限的默认归属问题。在传统治理安排中，决策权通过层级授权链条自然落到管理层，董事会只在重大事项上保留决定权；人工智能改变了这一前提——具备何种自主能力的系统能否进入某个场景、在何种约束下运行、何时必须停止，这些问题的答案并不天然存在于既有的授权文件之中。我们认为可以从以下方面着手：

确立董事会的保留事项清单。人工智能战略与重大部署的审查、重大风险敞口的评估、高风险应用的上线与退出，须经董事会或其下属专门委员会审议，不得以经营决策为由越过。委员会的组成应包含具备技术背景的外部专家，并保证独立董事占多数，其议事记录纳入董事会文件的正式归档范围。清单的价值不在于增加一道程序，而在于把原本默认属于管理层的决定权重新明确为董事会的决定事项，使授权关系在人工智能场景下仍然成立。

使叫停成为默认选项之一，而非仅在例外情形下才予考虑。控制手段的失效通常不出现在是否具备手段这一层，而出现在由谁在什么程序下决定是否启用这一层：安全配置的调低、监控的开启与关闭、运行的中断与重启，如果没有明确的责任主体与书面程序，就容易在能力建设或产品进度的压力下被默认放弃。可行的制度设计包括三项：将高风险评估与训练中安全配置的变更列为须经专门委员会书面批准的保留事项，并禁止由执行团队以进度为由自行决定；明确重大告警在时限内未获排除时的自动暂停效力，使暂停的发生不需要另行决策；对暂停与重启的权限作具名安排并要求留存决策记录，使事后可以查证由谁做出判断。与这一过程相配套，事故应急响应机制须明确触发条件、处置流程与对外报告路径，并把紧急停用权明确授予具体岗位而非抽象部门。

给数据处理的边界一个明确的授权主体。采用差分隐私、同态加密等技术可以在数据不出域、不可还原的前提下支持模型训练，从而在数据这一侧为授权提供技术支撑。但这些技术能够降低数据集中带来的泄露风险，却无法解决数据采集本身是否具备合法性基础的问题——技术手段能够保护已经合法采集的数据，不能为不当采集提供正当性。训练数据的来源、范围与使用目的，以及哪些个人信息不得进入模型的训练与优化，仍然需要由有权主体做出明确决定并留存记录。《关于银行业保险业人工智能安全开发应用的指导意见》（金发〔2026〕8 号）已针对银行业保险业划出了明确的数据红线，其他行业的企业虽不直接适用，但同一思路同样成立，把红线落到具体的授权决定之上，使每一次数据进入模型训练都有可指认的决定人。

（二）监督机制：让监督者够得着事实

监督环节要同时解决三件事，缺一件都会使监督落空：由谁监督，依据什么监督，以及能否取得做出判断所需的事实。董事自身难以完成对复杂模型的实质性审查，主体因此必须专业

下载日志已记录，仅供内部参考

化；即便主体到位，若没有可以对照的规则，审查只能停留在印象判断；而规则齐备之后，决定监督是否真实发生的，仍然是被监督者能否拒绝提供事实。我们认为可以从以下方面着手：

监督主体的设置宜分三层，分别对应决策、组织与技术三个层面。董事会专门委员会决定做什么、做到什么程度，其职责是审查人工智能战略与重大部署、评估重大风险敞口并定期听取报告。管理层设置首席人工智能伦理官，在管理序列中与首席技术官、首席风险官并列，并以独立预算保障工作条件，具体职责包括组织上线前的伦理影响评估、主持算法风险的定期评估、受理内部举报并在必要时向董事会直接报告。运营层面的算法审查小组承担技术性审查，包括偏见检测、安全测试、可解释性评估与留痕合规检查，其审查结论应当留档并可供内部审计复核。

专业能力缺口往往仍需外部力量来补充，使用外部专家需要附加三项约束：其一，专家意见应书面做出并归入档案，明确其依据的信息范围与所受的限制，避免依据管理层提供的材料成为意见的实际前提却未在文本中显现；其二，明确专家意见在决策链条中的地位，它是董事做出判断的输入而非替代品，采纳与否及采纳理由应当在议事记录中体现，判断责任不因采纳专家意见而转移；其三，专家的选聘与续聘不应由被监督的业务条线决定。这三项约束的作用是使外部意见可被检验，而不是使外部意见承担本应由董事承担的责任。

监督还须有可以对照的规则，否则审查只能停留在印象判断。需要把外部标准转译为内部规则，哪些业务场景允许部署具备自主执行能力的智能体，哪些必须限定在人工确认之后执行，哪些禁止使用。治理要求从重视风险变为达到某项指标方可上线，董事会与审查小组因此获得可核对的判断基准。与规则配套的是审查节奏的设计。按风险等级差异化安排：低风险应用采用自评与备案制，中风险应用采用集中审查，高风险应用适用完整的评估与审批程序，并明确各等级的判定标准与升级条件。上线前的伦理影响评估也应区分两种情形，纳入复核清单的高风险活动适用完整程序，其余适用简易程序，这一分级与《人工智能科技伦理审查与服务办法（试行）》确立的复核清单机制相对应；评估内容至少覆盖数据来源合法性、歧视风险评估、人工干预设计、退出与应急安排四个要素。定期报告则宜与现有风险管理报告合并，避免形成独立于既有治理体系之外的平行报告，从而弱化其被审议的严肃性。

主体与规则齐备之后，监督能否真实发生，取决于监督者能否取得事实。应在制度中明确三条：董事会专门委员会与伦理职能有权调阅任一在用模型的运行日志、推理痕迹与评估记录，且调阅范围不由被监督部门划定；为其配备独立于业务条线的技术资源，使其不必依赖被监督方提供的分析能力；对上述访问本身留痕，防止信息权被反向用于事后合理化。

（三）问责机制：把算法决策的责任锚定到具体个体

问责环节修复的困难不在组织设计而在概念前提。传统的归责路径假定存在一个做出决定的自然人或一个可辨认的决策集体，沿着谁决策谁负责回溯即可找到责任的落点；当行为由模型自主形成、沿着执行框架与环境反馈自行演化时，这一路径在关键的环节失效——既无法指认某个自然人做出了这个决定，也无法说某个集体在会议上批准了它。我们认为可以从以下方面着手：

把模型行为还原为若干个由人做出的前置决定，并在事前为这些决定配置责任主体。至少包括四类：是否部署，即把具备何种自主能力的系统投入何种场景；授予何种权限，即允许它访问哪些系统、调用哪些工具、在多大数据范围内运行；设置何种约束，即安全配置的水平、监控是否启用、告警如何处置；以及在异常发生时由谁决定是否停止。这四类决定无一例外由人做出，也都可以在事前具名。把它们固定为书面记录并在组织内部明确对应的岗位责任之后，即便行为本身由模型自主形成，责任仍然可以落在最初的授权与约束之上，而不是落在算法做出了决定这一句无法追责的描述上。这一处理与人机协同的原则是一致的：人工智

下载日志已记录，仅供内部参考

能在法律属性上始终是工具，最终决定权与责任归属必须落在人身上。

在多个岗位共同参与的情形下，还需要预先约定责任的分配顺序。上线审批、配置调整、监控启用分别由不同主体决定，损害发生后如果允许各自沿着我只批准了某一个环节退开，责任就会被稀释到无人承担。应在事前明确各环节的决策者就其决定范围内的事项承担首要责任，就超出其决定范围但应当上报的部分承担报告义务；业务变更导致风险等级提升而审批权限未随之调整时，不因流程未及而免除责任。这些规则不必等待外部认定标准的出现，企业可以在内部的问责办法中写明，而它们的存在本身就是事后追责得以进行的条件。

记录必须覆盖决策过程本身，而不只是运行结果。上线的审批记录、配置变更的时间与批准人、告警的处置过程、重启的授权，应当与模型的运行日志、推理痕迹一并留存，并存放于执行者无法自行修改的系统中。

风险提示

1、人工智能技术发展突破人工超智能；

2、企业公司治理对人工智能的挑战应对不足。

下载日志已记录，仅供内部参考

分析师声明

作者在中国证券业协会登记为证券投资咨询(分析师)，以勤勉的职业态度，独立、客观地出具本报告。作者保证：（i）本报告所采用的数据均来自合规渠道；（ii）本报告分析逻辑基于作者的职业理解，并清晰准确地反映了作者的研究观点；（iii）本报告结论不受任何第三方的授意或影响；（iv）不存在任何利益冲突；（v）英文版翻译若与中文版有所歧义，以中文版报告为准；特此声明。

投资评级说明

投资建议的评级标准		评级	说明
报告中投资建议所涉及的评级分为股票评级和行业评级（另有说明的除外）。评级标准为报告发布日后 6 到 12 个月内的相对市场表现，也即：以报告发布日后的 6 到 12 个月内的公司股价(或行业指数）相对同期相关证券市场代表性指数的涨跌幅作为基准。其中：A 股市场以沪深 300 指数为基准；香港市场以恒生指数为基准；美国市场以纳斯达克综合指数或标普 500 指数为基准。	股票评级	买入	预期个股相对同期相关证券市场代表性指数的回报在 20%及以上；
		增持	预期个股相对同期相关证券市场代表性指数的回报介于 5%～20%之间；
		持有	预期个股相对同期相关证券市场代表性指数的回报介于-10%～5%之间；
		卖出	预期个股相对同期相关证券市场代表性指数的回报在-10%及以下；
		无评级	预期对于个股未来 6 个月市场表现与基准指数相比无明确观点。
	行业评级	看好	预期行业整体回报高于基准指数整体水平 10%以上；
		中性	预期行业整体回报介于基准指数整体水平-10%～10%之间；
		看淡	预期行业整体回报低于基准指数整体水平-10%以下。

一般声明

五矿证券有限公司（以下简称“本公司”）具有中国证监会批复的证券投资咨询业务资格。本公司不会因接收人收到本报告即视其为客户，本报告仅在相关法律许可的情况下发放，并仅为提供信息而发放，概不构成任何广告。本报告的版权仅为本公司所有，未经本公司书面许可，任何机构和个人不得以任何形式对本研究报告的任何部分以任何方式制作任何形式的翻版、复制或再次分发给任何其他人。如引用须联络五矿证券研究所获得许可后，再注明出处为五矿证券研究所，且不得对本报告进行有悖原意的删节和修改。在刊载或者转发本证券研究报告或者摘要的同时，也应注明本报告的发布人和发布日期及提示使用证券研究报告的风险。若未经授权刊载或者转发本报告的，本公司将保留向其追究法律责任的权利。若本公司以外的其他机构（以下简称“该机构”）发送本报告，则由该机构独自为此发送行为负责。
本报告所载的资料、意见及推测仅反映本公司于发布本报告当日的判断，本报告所指的证券或投资标的的价格、价值及投资收入或将产生波动；在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告；本公司不保证本报告所含信息保持在最新状态。同时，本公司对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。
本报告的作者是基于独立、客观、公正和审慎的原则制作本研究报告。本报告的信息均来源于公开资料，本公司对这些信息的准确性和完整性不作任何保证，也不保证所包含信息和建议不发生任何变更。本公司已力求报告内容的客观、公正，但文中的观点、结论和建议仅供参考，不包含作者对证券价格涨跌或市场走势的确定性判断。在任何情况下，报告中的信息或意见不构成对任何人的投资建议，投资者据此做出的任何投资决策与本公司和作者无关。在任何情况下，本公司、本公司员工或者关联机构不承诺投资者一定获利，不与投资者分享投资收益，也不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。本公司及作者在自身所知情范围内，与本报告中所评价或推荐的证券不存在法律法规要求披露或采取限制、静默措施的利益冲突。
五矿证券版权所有。保留一切权利。

特别声明

在法律许可的情况下，五矿证券可能会持有本报告中提及公司所发行的证券并进行交易，也可能为这些公司提供或争取提供投资银行、财务顾问和金融产品等各种金融服务。因此，投资者应当考虑到五矿证券及/或其相关人员可能存在影响本报告观点客观性的潜在利益冲突，投资者请勿将本报告视为投资或其他决定的唯一参考依据。

联系我们

上海

地址：上海市浦东新区陆家嘴街道富城路 99 号震旦国际大厦 30 楼
邮编：200120

深圳

地址：深圳市南山区滨海大道 3165 号五矿金融大厦 23 层
邮编：518035

北京

地址：北京市东城区朝阳门北大街 3 号五矿广场 C 座 3F
邮编：100010

下载日志已记录，仅供内部参考

Analyst Certification

The research analyst is primarily responsible for the content of this report, in whole or in part. The analyst has the Securities Investment Advisory Certification granted by the Securities Association of China. Besides, the analyst independently and objectively issues this report holding a diligent attitude. We hereby declare that (1) all the data used herein is gathered from legitimate sources; (2) the research is based on analyst's professional understanding, and accurately reflects his/her views; (3) the analyst has not been placed under any undue influence or intervention from a third party in compiling this report; (4) there is no conflict of interest; (5) in case of ambiguity due to the translation of the report, the original version in Chinese shall prevail.

Investment Rating Definitions

The rating criteria of investment recommendations

The ratings contained herein are classified into company ratings and sector ratings (unless otherwise stated). The rating criteria is the relative market performance between 6 and 12 months after the report's date of issue, i.e. based on the range of rise and fall of the company's stock price (or industry index) compared to the benchmark index. Specifically, the CSI 300 Index is the benchmark index of the A-share market. The Hang Seng Index is the benchmark index of the HK market. The NASDAQ Composite Index or the S&P 500 Index is the benchmark index of the U.S. market.

	Ratings	Definitions
Company Ratings	BUY	Stock return is expected to outperform the benchmark index by more than 20%;
	ACCUMULATE	Stock relative performance is expected to range between 5% and 20%;
	HOLD	Stock relative performance is expected to range between -10% and 5%;
	SELL	Stock return is expected to underperform the benchmark index by more than 10%;
	NOT RATED	No clear view of the stock relative performance over the next 6 months.
Sector Ratings	POSITIVE	Overall sector return is expected to outperform the benchmark index by more than 10%;
	NEUTRAL	Overall sector expected relative performance ranges between -10% and 10%;
	CAUTIOUS	Overall sector return is expected to underperform the benchmark index by more than 10%.

General Disclaimer

Minmetals Securities Co., Ltd. (or "the company") is licensed to carry on securities investment advisory business by the China Securities Regulatory Commission. The Company will not deem any person as its client notwithstanding his/her receipt of this report. The report is issued only under permit of relevant laws and regulations, solely for the purpose of providing information. The report should not be used or considered as an offer or the solicitation of an offer to sell, buy or subscribe for securities or other financial instruments. The information presented in the report is under the copyright of the company. Without the written permission of the company, none of the institutions or individuals shall duplicate, copy, or redistribute any part of this report, in any form, to any other institutions or individuals. The party who quotes the report should contact the company directly to request permission, specify the source as Equity Research Department of Minmetals Securities, and should not make any change to the information in a manner contrary to the original intention. The party who re-publishes or forwards the research report or part of the report shall indicate the issuer, the date of issue, and the risk of using the report. Otherwise, the company will reserve its right to taking legal action. If any other institution (or "this institution") redistributes this report, this institution will be solely responsible for its redistribution. The information, opinions, and inferences herein only reflect the judgment of the company on the date of issue. Prices, values as well as the returns of securities or the underlying assets herein may fluctuate. At different periods, the company may issue reports with inconsistent information, opinions, and inferences, and does not guarantee the information contained herein is kept up to date. Meanwhile, the information contained herein is subject to change without any prior notice. Investors should pay attention to the updates or modifications. The analyst wrote the report based on principles of independence, objectivity, fairness, and prudence. Information contained herein was obtained from publicly available sources. However, the company makes no warranty of accuracy or completeness of information, and does not guarantee the information and recommendations contained do not change. The company strives to be objective and fair in the report's content. However, opinions, conclusions, and recommendations herein are only for reference, and do not contain any certain judgments about the changes in the stock price or the market. Under no circumstance shall the information contained or opinions expressed herein form investment recommendations to anyone. The company or analysts have no responsibility for any investment decision based on this report. Neither the company, nor its employees, or affiliates shall guarantee any certain return, share any profits with investors, and be liable to any investors for any losses caused by use of the content herein. The company and its analysts, to the extent of their awareness, have no conflict of interest which is required to be disclosed, or taken restrictive or silent measures by the laws with the stock evaluated or recommended in this report.

Minmetals Securities Co. Ltd. 2019. All rights reserved.

Special Disclaimer

Permitted by laws, Minmetals Securities Co., Ltd. may hold and trade the securities of companies mentioned herein, and may provide or seek to provide investment banking, financial consulting, financial products, and other financial services for these companies. Therefore, investors should be aware that Minmetals Securities Co., Ltd. or other related parties may have potential conflicts of interest which may affect the objectivity of the report. Investors should not make investment decisions solely based on this report.

Contact us

Shanghai

Address: 30/F, Zhendan International Building, No.99 Fucheng Road, Lujiazui Street, Pudong New District, Shanghai
Postcode: 200120

Shenzhen

Address: 23F, Minmetals Financial Center, 3165 Binhai Avenue, Nanshan District, Shenzhen
Postcode: 518035

Beijing

Address: 3/F, Tower C, Minmetals Plaza, No.3 ChaoyangmenNorth Street, Dongcheng District, Beijing
Postcode: 100010