人工智能侵权立法中的归责逻辑冲突与矫正*

徐榕岭
（互联网法治研究院（杭州），浙江　杭州　310000）

摘　要：人工智能技术的发展对侵权立法形成挑战，基于人工智能风险分级的侵权归责逻辑与传统归责逻辑的冲突凸显。理论上，基于人工智能风险分级的侵权归责逻辑源于欧盟的人工智能风险分级治理方案，但其冲击了既有民事侵权归责结构，使各方主体的责任承担处于不确定状态。将技术风险凌驾于权利保护客体差异之上，本质上是一种混淆行政法与侵权法适用边界的错误做法。人工智能技术风险具有损害严重、逃逸诉讼可能性高、信息成本大等特征，难以在权利保护规则下得到有效规制。此时，回归基于权利客体差异的归责逻辑确有必要。具体而言，应先明确行政法与侵权法的分工，以“公共利益”为二者界分标准，必要时可通过原则性规定有限地接纳新型风险的规制任务，以最小制度成本实现技术风险的高效治理。

关键词：人工智能；风险分级治理；侵权立法；民行衔接
中图分类号：D923；D922.17　　**文献标识码**：A　　**DOI**：10.19358/j.issn.2097-1788.2025.12.011
引用格式：徐榕岭. 人工智能侵权立法中的归责逻辑冲突与矫正［J］. 网络安全与数据治理，2025，44(12)：75-82.

Conflicts and corrections in the attribution logic of tort liability legislation for artificial intelligence

Xu Rongling
(Internet Governance Institute (Hangzhou), Hangzhou 310000, China)

Abstract: The development of artificial intelligence (AI) technology poses significant challenges to tort liability legislation, particularly through the conflict between traditional attribution logic and the emerging risk-tiered liability framework. Theoretically, the risk-tiered attribution logic originates from the European Union's risk-based governance scheme for AI. However, it disrupts existing civil tort liability structures, leading to uncertainty in liability allocation among stakeholders. By prioritizing technological risks over distinctions in protected legal interests, this logic inherently conflates the jurisdictional boundaries between administrative law and tort law. AI-related risks—characterized by severe damages, high likelihood of evading litigation, and substantial informational costs—are difficult to regulate effectively under conventional rights-protection rules. Consequently, it is imperative to revert to an attribution logic grounded in distinctions among legal interests. Concretely, a clear division of responsibilities between administrative and tort law must be established, with "public interest" serving as the demarcation criterion. When necessary, principled legal provisions may be adopted to address emerging risks in a limited manner, thereby achieving efficient governance of technological risks with minimal institutional costs.

Key words: artificial intelligence; risk-based regulation; tort liability legislation; civil-administrative interface

0　引言

随着人工智能（Artificial Intelligence，AI）技术的兴起，“风险”逐渐成为归责核心。经济分析中，“风险”是指决策中的度量不确定性[1]。那么“风险”作为归责中心即可能发生的不利结果应被视为一种成本并纳入制度管理与市场调节中。具体到AI技术的法律规制层面，“风险”的概念内涵已逐渐被“技术损害”填充，即AI技术可能引发的损害是一种应当被控制的社会成本。由此，欧盟提出了AI风险分级治理方案，借此控制AI技术对社会安全、基本权利等的有害影响。但AI风险分级治理方案并不停留在行政法层面，也对侵权法的立法与解

*基金项目：2025年度中共杭州市委政法委、杭州市法学会法学研究课题“人工智能相关新质生产力的司法保障研究”

释产生深刻影响。例如，将 AI 技术风险直接视为新的侵权损害类型[2]、将依 AI 系统的风险等级来调整其适用的因果关系认定方式以及将高风险 AI 系统归入严格责任制度等。尽管将 AI 风险分级治理方案引入侵权立法有一定合理性，但这一做法也将与侵权归责逻辑相互冲突，引发不必要的制度成本。

当前，我国《人工智能法》立法尚处热议阶段，不同版本的专家建议稿都显示出对欧盟 AI 风险分级治理方案的偏向[3]。但其对于侵权归责逻辑的设计仍十分模糊，没有充分回应 AI 风险分级治理方案引入后的民行衔接难题。由此，本文将在考虑行政法与侵权法适用边界的基础上，提出能够妥善缓解新型 AI 技术风险救济难题，又能兼顾法律体系化的新方案。具体而言，本文第一部分将详细说明“基于 AI 风险分级的侵权归责逻辑”的概念、形成以及弊端，揭示 AI 风险分级治理方案介入后的一系列民事侵权体系冲击；第二部分将纠正隐藏于“基于 AI 风险分级的侵权归责逻辑”下的认识误区，回归“基于权利客体差异的侵权归责逻辑”；第三部分将论述“基于权利客体差异的侵权归责逻辑”的内部构造、界定的行政法与侵权法基本适用边界，同时说明用于接纳新型 AI 技术风险的原则性规定的适用方法，以纾解人工智能侵权立法中的归责逻辑冲突。

1　基于 AI 风险分级的侵权归责逻辑新议与矛盾呈现

1.1　基于 AI 风险分级的侵权归责逻辑之形成

所谓基于 AI 风险分级的侵权归责逻辑，即根据行为对象或使用场景的不同风险程度预先对 AI 技术进行风险分级，再依其所属风险等级来匹配不同类型因果关系、责任原则、注意义务的归责方式。其源于欧盟的 AI 风险分级治理方案，是原于公法领域运行的风险识别与控制机制向民事侵权体系延伸的结果[4]。2020 年《欧洲议会关于人工智能民事责任制度的决议》（The 2020 AI Liability Resolution）中，第 14 条明确提出应当认识到 AI 系统类型是确定侵权责任的关键因素，应将高风险 AI 系统归入严格责任制度，而非沿用过错责任[5]。责任原则的改变即为部分侵权归责结构的改变。在欧盟 AI 风险分级治理方案逐渐渗透到我国 AI 治理体系过程中，亦出现了类似观点[6]。例如，禁用型 AI 系统应适用严格责任原则，高风险型和低风险型 AI 系统则应适用过错责任原则[7]。这一做法几乎彻底推翻了原本基于权利客体差异的侵权归责逻辑，无疑会使利益相关者的侵权责任承担陷入混乱状态中。

另一包含基于 AI 风险分级的侵权归责逻辑的突出观点则认为，AI 技术会引发我国立法尚未承认的新型损害，如数据信息损害[8]。直接在民事侵权体系下承认 AI 技术风险，也是一种直接引入 AI 风险分级治理方案的方式。其背后的逻辑大致可以理解为：先承认 AI 技术可能造成不同等级的风险，风险级别不与既有权利保护客体完全对应；随后，选择具有实质性危害的 AI 技术风险纳入民事权利保护范畴之中，以达到以个人诉讼或集体诉讼方式来缓解 AI 技术风险危害，弥补权利人的损失。即便不作明示，欧盟提出的责任原则转变方案也已经隐含了这一观点。其原因在于，依 AI 技术风险等级改变而改变责任原则的前提是承认新型 AI 技术风险的存在，否则能够救济的 AI 技术风险非常有限。因此，可以将上述对照 AI 风险分级治理方案建设民事权利体系的做法视为基于 AI 风险分级的侵权归责逻辑的直接引入。仅从 AI 技术风险治理的角度看，直接引入未必效果不佳，但仍然存在精细化构建的空间。

不过，在各国已建立相对稳定的民事侵权体系之现实背景下，直接引入 AI 风险分级治理方案不太现实。例如，欧盟《人工智能责任指令（提案）》（The Artificial Intelligence Liability Directive，AILD）提出，应于比例原则指导下针对性协调举证规则，而非统一民事责任规则[9]。尽管 AILD 没有明确将 AI 风险分级治理方案排除在民事侵权体系之外，但这仍然能体现出欧盟对于高昂转换成本或体系稳定性的担忧。那么基于 AI 风险分级的侵权归责逻辑是否仅为一种不切实际的理论推理？答案是否定的，因为 AI 风险分级治理方案仍可借侵权归责结构的变革渗入，亦即基于 AI 风险分级的侵权归责逻辑的间接引入。

第一条路径中，AI 风险分级治理方案是通过该举证规则的改变完成渗透的。区别于欧盟《人工智能法案》（The Artificial Intelligence Act）的前端规制，AILD 在后端补充了部分举证规则[10]，根本目的是降低权利人的胜诉难度。例如，AILD 第 3 条规定了在原告能够提出合理理由的情况下，可以要求法院命令被告披露与案涉 AI 系统相关的证据。尽管欧盟声称 AILD 只是有助于排除错误被告，节省时间、费用和法院案件负担，但 AILD 仍实质性改变了诉讼双方举证责任分配情况。适用 AILD 第 3 条后，被定义为高风险的 AI 系统须向原告披露相关证据，那么原本由权利人承担的举证责任便很大程度上转移到被告一方。此时，AILD 翻转了技术方适用的责任原则，使举证责任倒置，充分体现了 AI 风险分级治理方案对于侵权归责结构的重塑能力。如果我国立法者后续接纳 AI 风险分级治理方案，不断深化披露义务，那么这一转变也将会在我国发生。

第二条路径中，AI 风险分级治理方案则是通过公法

义务的私法化完成渗透的。即行为人没有完成相应风险等级的行政法义务，就推定其存在过失。实践中，行政法义务向侵权法过渡的做法十分常见，如医疗损害责任认定与行政法规定的诊疗护理规范挂钩。但 AI 领域存在巨大的信息成本，尤其体现在人类对 AI 技术决策机制的理解困难上[11]。这将直接导致行政层面合理的、符合实际情况的行为人义务的设计困难。一旦行为人需要完成的行政法层面的义务要求过高，加之法律允许“违法推定过失”，那么包括 AI 服务提供者等责任主体都必须承担更重负担。例如，我国《生成式人工智能服务管理办法》要求生成式 AI 服务提供者笼统承担反内容歧视和反虚假生成、训练数据合规等注意水平要求较高的义务，但生成式 AI 服务提供者实际无法在合理成本范围内完成上述任务。那么程度较高的行为义务便会实质性改变 AI 服务提供者适用的责任类型，趋向严格责任。这一过程亦体现了 AI 风险分级治理方案对于侵权归责结构的重塑能力。

1.2 基于 AI 风险分级的侵权归责逻辑之弊端

结合 AI 风险分级治理方案入侵民事侵权归责结构的路径分析，可以发现基于 AI 风险分级的侵权归责逻辑扰乱了民事侵权归责结构。因果关系认定亦是如此。AILD 第 4 条允许推定高风险 AI 系统与损害之间因果关系后，AI 风险分级治理方案也重塑了因果关系认定要件。无论是因果关系、举证责任或过错认定，都属于 AI 服务提供者、用户、权利人等利益相关者之间责任分配问题解决的前置障碍。不过，由于涉 AI 技术侵权案件中因果关系认定逐渐式微，整体性利益衡量正逐步取代要件式认定思路，因而 AI 风险分级治理方案对因果关系认定的影响不及其他要件明显。但整体来看，如果不能排除 AI 风险分级治理方案对利益相关者所适用的侵权归责结构的不利影响，那么责任分配的问题仍然难以展开。

未引入 AI 风险分级治理方案之前，民事侵权归责结构通常建立在权利保护客体的差异上。例如，著作权保护作品利益，而作品利益主要经由“使用”实现，那么著作权侵权归责结构的选择就必须依作品的“使用”特征展开。在多数作品“使用”都无法避免地存在过错时，著作权自然会趋向严格责任，加重行为人的举证义务和注意义务，尤其是在署名权认定过程中[12]。又例如，人格权保护的是人格利益，但大部分人格利益损害存在一定量化障碍[13]，导致其利益使用与利益损害之间的边界相对模糊。此时，如果选择严格责任，则可能损害言论自由。可见，不同权利保护客体匹配不同侵权归责结构，这是具有效率的做法。但引入 AI 风险分级治理方案之后，权利保护客体的差异就会被逐渐淡化，技术风险差异则会被逐步放大并成为责任原则选择的关键。放大技术风险的直接后果是损害体系将被重新定义。例如，AI 生成内容引发的人格利益损害会分化为虚假信息风险、歧视信息风险等，但这与原本的权利保护客体划分体系并不匹配。这便是忽视权利保护客体差异的结果，势必与民事侵权归责体系冲突。暂不论 AI 风险分级治理方案本身是否可靠，仅制度转换成本就已远超可承受范围，即行政法层面的改革引发了民事层面的制度负担。

其实，欧盟 AI 风险分级治理方案包含着明确的政策目的，即严厉管制包括 AI 技术在内的新兴技术。所谓由 AI 服务提供者或 AI 技术本身带来的“高风险”，只是用于支撑加强监管的一种正当性解释。这与美国宽松的技术发展环境正相反，即美国采用分散的“联邦弹性监管＋州级场景立法”模式[14]，在避免风险分级的同时，回应了 AI 治理需求。追根究底，这是由地区或国家技术发展水平决定的。对于欧盟而言，只有尽快促成“欧洲方案”的落实，才能为自己在全球 AI 规范领域成为权威并掌握话语权。但 AI 风险分级治理方案显然还存在很多漏洞，导致部分风险对应的技术方责任不升反降。这不仅没有完全达到严格管控换取本土技术发展空间的目的，也使得技术方的侵权责任将变得更为模糊。近期，欧盟撤回 AILD 的做法就是其察觉基于 AI 风险分级的侵权归责逻辑弊端的最好例证[15]。同理，在我国已经签署《布莱切利宣言》并加入 AI 国际治理，而 AI 风险分级治理方案借鉴呼声又逐渐高涨的当下[16]，归责逻辑问题的澄清变得十分关键。

2 基于 AI 风险分级的侵权归责逻辑的认识偏差纠正

2.1 用于分析与纠正认识偏差的 Shavell 框架

明确基于 AI 风险分级的侵权归责逻辑会与侵权归责结构相互冲突后，还需找出冲突背后的根本原因，以便恢复适宜的涉 AI 技术侵权之归责环境。本文认为，新侵权归责逻辑的争议核心是行政法与侵权法的平衡问题，或管制规则与责任规则的界分问题。具体而言，基于 AI 风险分级的侵权归责逻辑本质是希望将部分行政法层面的规则腾挪至侵权法之中，以减轻事后追责的压力，毕竟部分损害没有被既定的权益框架所接纳。但其没有认识到二者之间存在明确的界分标准。例如，过于分散的不利后果会导致单个受害者难以起诉，需要政府统一解决等。行政法与侵权法的设计不平衡直接导致了事前管制与事后追责混同，即为基于 AI 风险分级的侵权归责逻辑下隐含的归责认识偏差。对此，本文将主要借“Shavell 框架”找出每一认识偏差并加以改正，以期彻底解决归责逻辑问题。

所谓“Shavell 框架”，即一系列比较责任规则与管制规则适用利弊的标准，由 Steven Shavell 提出[17]。其中包含四个具体评价标准，即支付能力差异、逃避诉讼可能性差异、信息优势差异以及行政成本差异。整体来看，Shavell 框架是届分责任规则与管制规则的经济分析方法，可以弥补现阶段教义学分析的不足。就责任规则与管制规则的互动关系而言，既往研究多以现行法秩序作前提，通过案例分析找出“标准观点”，从而寻求现行法秩序下的妥当方案并找到规范依据[18]。但教义学分析缺乏对制度选择背后经济效率的系统考察，通常难以解释不同风险类型与成本条件下责任规则与管制规则的最优配置。相反，Shavell 框架通过引入不同成本维度，为二者的制度边界和组合使用提供了更精确的判断工具，能为新兴技术领域的规则设计提供理论支撑。因此，利用 Shavell 框架解决 AI 侵权立法中的归责逻辑选择问题具有正当性。

2.2 Shavell 框架下新侵权归责逻辑的偏差纠正

2.2.1 AI 技术风险不必构成合格民事损害

基于 AI 风险分级的侵权归责逻辑强调以“风险”替换或充实“权益”这一传统概念。此时，许多未被民事权利保护规则认可的利益也被纳入保护行列，如威胁教育或职业机会平等的社会机会风险、威胁基本自由的社会歧视风险等[19]。但“权益”泛化本身就是需谨慎对待的问题，即需要考虑所列 AI 技术风险是否值得被法律规制。适用基于 AI 风险分级的侵权归责逻辑，使 AI 服务提供者或用户承担不必要的风险制造责任，会损伤技术发展激励。诚然，也存在司法者直接认定新权益或新损害的做法，但前提是某一新权益确实经过实质性分析和权衡，法院能够确认不保护这一新权益将会引发负外部性[20]。理论上，基于 AI 风险分级的侵权归责逻辑下的新利益也应当经过类似的分析，才能被民事权利保护规则囊括。但 AI 风险分级治理方案中的新利益的变化速度非常快，不具有相对固定性，甚至难以通过行政规则的检验，不适合单列为民事权利保护规则中的新法益。例如，南京市中级人民法院“Cookie 隐私”案中，法院拒绝承认基于隐私权侵权行为会严重损害“精神安宁”利益。这充分说明并非所有新风险都应被认定为是合格民事损害，亦侧面印证基于 AI 风险分级的侵权归责逻辑没有认识到财产规则的确立对侵权归责有重要影响。

另一种情况是行政法罗列的 AI 技术风险属于应被法律救济的损害，但没有必要放置于侵权法下进行规制。这很大程度是由责任主体的支付能力过低导致的。Shavell 框架下，决定事前管制和事后追责的关键点之一是行为人的责任能力（Ability to Pay for Harm）。即如果行为人具备足够的赔偿能力，能够覆盖其可能造成的损害，那么可以主要依靠事后侵权责任救济；如果损害规模巨大，超过行为人的赔偿能力，那么应优先采用事前管制。这一标准亦可以引申为，如果损害发生带来的社会成本无法填补，那么应做充分的准备预防其发生。回到 AI 技术侵权场景中，可以发现法律上的禁用风险或高风险都具有很强烈的基本权利特征，属于人的根本性权利的细化。一旦遭遇侵害，严重性必然较一般权利损害更大。由此，如果 AI 技术风险的确难以补救，就应当通过限制交易条件的方式来尽可能限制损害发生，将风险处理前置，而非等损害发生再矫正。可见，基于 AI 风险分级的侵权归责逻辑中存在对 AI 技术风险性质、体量与技术方责任能力的认识误区。

2.2.2 AI 技术风险存在极大逃逸可能性

沿着 AI 技术风险的特性分析，可以发现基于 AI 风险分级的侵权归责逻辑还忽视了 Shavell 框架下的诉讼可及性问题。诉讼可及性与逃避诉讼可能性同理，亦为事前管制与事后追责的判断标准之一。其含义是如果权利人能更方便地识别或起诉行为人，那么事后追责能给予有效威慑；如果行为人能轻易逃避诉讼，那么应当选择事前管制。在 AI 技术侵权场景中，多数 AI 技术风险具有非常强的分散性。例如，威胁教育或职业机会平等的社会机会风险，可以被拆解出一部分隶属于民事权利保护规则，即美国《民权法案》（Civil Rights Act of 1964）具有私法属性的第七章所保护的公平就业利益，而其他部分又可以拆解到其他民事法律中救济或直接不被侵权法涵盖。但值得注意的是，可以被拆解的 AI 技术风险占比很小。大部分 AI 风险分级治理方案所描述的风险都不具备民事追责的可能，无论是具体权利方面还是主体资格方面。此时，行为人逃避诉讼可能性非常高，几乎没有适格主体能够直接提起与 AI 技术风险有关的诉讼。

承认 AI 技术风险存在较高逃避诉讼可能性后，基本可知 AI 技术本身所引发的风险仍应通过事前管制路径解决，即直接由行政机关进行控制和处理。当行政机关成为风险的主要控制人，那么 AI 服务提供者便不构成风险的唯一控制者。亦没有必要启动基于 AI 风险分级的侵权归责逻辑，再沿此形成较高义务要求来威慑 AI 服务提供者或倒逼其提高自身的风险防范水平。不过，反对者可能认为行政机关监督 AI 技术风险是否转变为实际损害的成本十分高昂，需耗费大量人力、物力，不及权利人直接通过诉讼精准打击有效。但上述观点忽视了民事权利保护规则中的现实约束条件，除非权利保护客体被法律扩宽，否则权利人甚至无法提起具有“实际损害”的诉讼。而改变民事权利保护规则亦需付出制度转变成本。

除非制度转变成本小于事后追责能够弥补的少量权利保护收益，否则没有改变的必要性。

2.2.3 AI 技术风险的规制要求较高信息成本

在支付能力差异与逃避诉讼可能性差异外，Shavell 框架还有信息优势差异以及行政成本差异两个因素。其含义是如果行为人比行政机构更了解其行为风险及防范方式，那么应采用事后追责模式，促使行为人根据信息优势来改进行为以防止风险落实；以及如果普遍性事前管制的行动成本过高，那么也应当倾向使用事后追责，否则会降低整体效率。不过，在 AI 技术侵权场景中，信息掌握能力与行政成本基本可以合并为技术风险知悉程度一起讨论。行政机关对 AI 服务提供者的侵权行为进行管制，也须以掌握 AI 技术信息为前提。例如，行政机关规范虚假新闻、深度伪造视频时，需要先掌握内容生成机制、输出通道、内容后处理流程、嵌入式水印技术原理、标识可验证性以及对系统性能的影响等技术信息，才能准确设定内容标识的技术标准。

碍于相关技术信息的秘密性，通常只有 AI 服务提供者能够构成最低信息成本一方。此时，Shavell 框架可能会导出事后追责更具效率的结论，因为行为人基于自我保护会主动利用信息优势来预防风险。但 Shavell 框架的一个关键前提是：行政机关不主动收集信息也不提前干预技术发展。这在现实中几乎不可能存在，因为国家同时存在权力和社会稳定的需求[21]，行政机关不可能放弃安全保障要求。一旦要求行为人进行定向披露或报备技术细节和潜在风险，那么行政机关便不再是信息不足的一方，反而具有合理、精准地制定事前管制计划的能力。此时，适度的事前管制便比纯粹依赖事后追责更加高效。如果一定要选择基于 AI 风险分级的侵权归责逻辑，那么技术信息的披露势必会与商业秘密的保护相冲突。在权利人没有直接获取 AI 服务提供者商业秘密的可能的背景下，扩大信息披露范围会引发更为严重的商业秘密泄露威胁。因此，行政机关与 AI 技术风险规制的高信息量要求相契合是双向推导出的结论，使行政机关同时承担技术信息审查和风险事前管制是更好选择。

3 权利保护客体差异上的侵权归责逻辑回归与完善

3.1 基于权利客体差异的侵权归责逻辑之必要性

所谓基于权利客体差异的侵权归责逻辑，即根据权利客体的性质与特征来确定归责原则等要件的归责方式。实效层面上，AI 侵权归责逻辑决定了各方承担侵权责任的多少。如果适用基于 AI 风险分级的侵权归责逻辑，那么在从风险到权益的重组过程中，不同权益对应的归责方式就会发生改变。出于平衡专有利益与公共利益的目的，须恢复基于权利客体差异的侵权归责逻辑作为归责前提。

第一，基于权利客体差异的侵权归责逻辑可以通过划定客体边界保障公私利益平衡。法律问题的本质是资源分配问题，而基于权利客体差异的侵权归责逻辑是保障资源分配符合广泛正义标准的手段之一，可以保证公私利益平衡。例如，基于权利客体差异划定的保护规则可以通过权利实现与权利限制规则的交叉设计，如基于作品特征而设计的复制、演绎、传播三类有限的基本权项，来避免控制范围过度扩张而影响公众对作品的接近，保证著作权人与作品使用者之间的利益分配均衡。即便是人格权一类排他性更为强烈的权利，也需要留出一定空白，如让渡于言论自由。遵守基于权利客体差异的侵权归责逻辑，可以避免从风险到利益的重组过程侵占使用者的利益实现空间，进而避免 AI 服务提供者被迫承担更重责任。

第二，基于权利客体差异的侵权归责逻辑可以通过确定举证责任、行为义务等具体归责结构保障公私利益平衡。强化权利客体差异的一个突出表现即责任类型随客体保护的变化而变化。一旦权利客体保护随其所处背景发生变化，那么责任类型也会随之变化。例如，个人信息权益保护进程中，由于信息处理具有特殊性，会引发权利人举证困难，于是原本属于人格利益分支的个人信息权益亦逐渐趋向过错推定责任或无过错责任[22]。这实际是扩张权利客体边界保障公私利益平衡的实例，能够保障不同损害分别适用不同的侵权归责方式。不难发现，即便是向外扩张责任范畴，基于权利客体差异的侵权归责逻辑仍然保持着相当程度的谦抑性，并未“包揽”AI 技术风险，仍为行为人留下了充分利益实现空间。

3.2 基于权利客体差异的侵权归责逻辑之内部构造

3.2.1 以行政法与侵权法的合理分工为归责前提

确定基于权利客体差异的侵权归责逻辑具有必要后，还需进一步完善其适用方式，使其更好地适应现实发展需要，尤其是存在行政法与侵权法协同规制需要时。一方面，尽管很多 AI 技术风险不构成合格民事损害或没有必要作为民事损害加以救济，但仍存在特定的需侵权法配合规制的情况，需以行政法与侵权法的合理分工为前提。立足侵权法适用立场，上述行政法与侵权法的协同规制可以解释为行政法对侵权法的补充作用。即各方对部分 AI 技术风险有充分认识与救济经验后，Shavell 框架下成本考量会发生改变，这些 AI 技术风险可以转变为相对固定的损害类型，适用侵权法进行规制更具效率。另一方面，从“基于 AI 风险分级的侵权归责逻辑”转换到“基于权利客体差异的侵权归责逻辑”，最为关键的步骤

便是排除行政法对侵权法适用的不当干预。此时，合理届分行政法与侵权法则更为必要，否则不能得知何种 AI 技术风险应当划归侵权法处理。因此，只有在厘清行政法与侵权法适用边界的基础上，才能开展二者的协同适用工作。

具体而言，行政法与侵权法的界限需借管制规则与财产规则、责任规则之间的关联关系来认定。理论上，自由市场通常能够通过价格机制来有效配置资源，而市场失灵则需要政府干预[23-24]。上述用于修正“市场调节不能”的手段即管制规则，是与财产规则或责任规则对立的调控手段。Shavell 框架下的事前管制即对应管制规则，如欧盟《人工智能法案》或《通用数据保护条例》中的行政规则；而事后追责则对应财产规则或责任规则，即侵权法中的产权规则与救济规则。基于权利客体差异的侵权归责逻辑的建立，要求三者进行良性互动。从交易成本的角度切入，管制规则的适用需要满足“财产规则部分失效”以及“责任规则成本过高”两个条件：首先，财产规则部分失效是进入责任规则与管制规则适用的前提。财产规则部分失效，即财产规则无法有效保障权利的排他性。例如，著作权人无法阻止 AI 服务提供者利用其作品进行数据训练，亦无法以合理交易成本与使用者缔约。当财产规则部分失效时，法律或允许权利人利用责任规则进行事后追责，或允许利用管制规则进行事前管制。其次，结合 Shavell 框架的使用情况，可以发现管制规则与责任规则的区分标准是总成本的多少。假设事前管制与事后追责能够获得的收益相等，即 Shavell 框架下行为人的责任能力相同时，哪一规制手段的总成本最低，则应被采纳。此处的总成本就包括支付可能性成本、逃避诉讼可能性成本、信息成本以及行政成本。

回到司法实践中，规制成本与收益之间的差值通常是以“公共利益”来表述。例如，AI 技术存在威胁公众平等、自由等基本权利的极大可能，进而损害“公共利益”，因而需要采取披露、报备或审计等措施。此时，明确“公共利益”的定义是区隔传统民事权利客体与 AI 技术风险的重要手段，可以科学地使某一风险或损害进入对应控制途径之中。同时，这也能避免 AI 技术风险已经超出责任规则承受范围，但仍没有管制规则的介入的情况。结合 Shavell 框架下的分析，可得构成“公共利益”必须具备：第一，某一 AI 技术风险不能在既有权利客体框架中找到完全对应项；第二，某一 AI 技术风险不能找到确定的权利人，或不存在对应的集体诉讼、公益诉讼途径；第三，AI 技术风险一旦落实，会引发超出个体层面的系统性后果，导致广泛且不可逆的社会损害；第四，现有法律规范和市场机制无法在可接受成本范围内实现对该 AI 技术风险的有效规制。上述条件的满足是 AI 技术风险进入事前管制的阀门。

3.2.2 以原则性条款的有限拓展适用为归责补充

除管制规则的补充外，基于权利客体差异的侵权归责逻辑的适用，还需配合民事权利保护规则的弹性塑造。当事前管制与事后追责能够获得的收益相等，且事后追责成本变得更低时，民事权利保护规则须灵活接受新法益保护需要。具体可考虑利用民事权利保护规则内部的原则性条款承接具体法益保护需要，直至其固化为一种新的民事权益。实践中，中华人民共和国《民法典》中的公序良俗条款、《反不正当竞争法》中的一般条款常被用于保护尚未权利化的法益。这充分说明原则性条款具有相当的适用基础。在排除侵犯“公共利益”可能性的基础上，允许原则性条款承接新的法益保护需要是法律灵活性的表现。经一段时间的检验后，亦可考虑固化权利。在 AI 技术风险层出不穷的背景下，原则性条款的有限适用以及权利固化过程，恰好能够匹配新型 AI 技术风险的生成特征，成为基于权利客体差异的侵权归责逻辑中“动态性”的主要来源。诚然，是否保护某一权益或是否规制某一行为仍然要以行动的社会福利最大化为标准，如果权益保护本身不能增进社会福利，那么无论是事前管制还是事后追责都不具有正当性。

然而，这一做法亦存在两方面限制：第一，新法益保护须避免与既有权利相互重复。例如，适用《反不正当竞争法》一般条款时，应禁止新法益保护与知识产权保护相互重复。这是为了确保原则性规定的适用不扩张既有权利规则的保护范围，保护既有利益平衡机制[25]。第二，新法益保护须严格遵守既有权利保护范式，尤其是在具体的侵权归责方式选择上。民事权利保护长期基于权利客体差异展开，那么新法益保护亦应当遵循既有思路来确定侵权归责方式，满足体系化要求。亦即仅依据侵权法内部要求对具体 AI 技术风险或损害的归责方式进行微调。首先，只有遭遇显著证据偏在时，可以考虑适用过错推定原则来缓解权利人举证成本过高的问题，如个人信息权益保护。其次，过错责任与严格责任之间的界限不可随意破除。同源或类似权利应适用相同的责任原则，责任类型的改变与责任承担程度勾挂，须通过实质审查才考虑变更。最后，过错认定中的注意义务设定需通过成本—效用分析法的检验[26]，避免过低或过高的行政法义务被无条件接收为过错判定标准。这也是锁定侵权责任入口的重要一环。

此处仍以虚假信息的规制为例，对基于权利客体差异的侵权归责逻辑的适用方法进行说明：当与事实不符、具有误导性或欺骗性信息生成时，需要先判断其是否属

于“公共利益”范畴。依据前文归纳的“公共利益”四个要件，纯粹的虚假信息并不与特定损害类型相互匹配，亦没有对应的权利人，除非虚假信息涉及人格利益。加之虚假信息具有严重负外部性，可能误导公众、破坏信息生态。此时，赋权行政机关对其进行整治便可以整体性降低个体追责成本，提前预防 AI 技术风险落实。而行政机关所需成本亦在可控范围之内，可以配合公众的监督和举报制度。例如，行政机关可以要求生成式 AI 服务提供者建立通知—删除制度，与美国近期通过的《删除法案》（Take It Down Act）构造类似[27]。当公众发现存在虚假信息时，允许先通知 AI 服务提供者删除内容。随后，如果反复出现或虚假信息更为严重，则可以统一交由行政机关进行处理，缓解无对应权利人便无人监管的问题，维护“公共利益”。进一步说，如果未来客观条件发生变化，虚假信息能够成为独立的民事权利损害类型，那么需要适用其同源或类似权利的侵权归责方式，除非具有特定事实足以改变此种推定。至此，基于权利客体差异的侵权归责逻辑才真正完成适用。

4 结论

在欧盟的 AI 风险分级治理方案被普遍接受和模仿后，出现了依行为对象或使用场景所属风险等级来匹配不同的归责方式的做法，即基于风险分级的 AI 服务提供者侵权归责逻辑。但此种侵权归责逻辑会冲击侵权归责结构，造成侵权责任认定混乱。其实，基于 AI 风险分级的侵权归责逻辑本质是事前管制与事后追责的界分混乱。归责逻辑混同会抵触既有侵权归责结构并影响治理效果。此时，需要认识到基于权利客体差异的侵权归责逻辑必要性和可行性。理论上，本文提出的是一种法律解释方案，不一定依赖于独立的《人工智能法》的建立与推出。但如果我国《人工智能法》的立法进程加快，那么可以考虑在其中合理界定“公共利益”这一概念，使其成为事前管制与事后规制的有效区分点。随后，有限适用《民法典》或部门法中原则性规定，接纳确有必要为民事权利所保护的对象。

参考文献

[1] 弗兰克·H·奈特. 风险、不确定性与利润 [M]. 安佳，译. 北京：商务印书馆，2017.

[2] SOLOVE D J, CITRON D K. Risk and anxiety: a theory of data-breach harms [J]. Texas Law Review, 2017, 96: 737 – 786.

[3] 人工智能法（学者建议稿） [EB/OL]. (2024 – 03 – 16) [2025 – 07 – 08]. https://mp.weixin.qq.com/s/wZCWNyUbMOi1-v1LBe6IZQ.

[4] CHAMBERLAIN J. The risk-based approach of the European Union's proposed artificial intelligence regulation: some comments from a tort law perspective [J]. European Journal of Risk Regulation, 2023, 14 (1): 1 – 13.

[5] European Parliament. Resolution of 20 October 2020 on a civil liability regime for artificial intelligence (2020/2014 (INL)) [EB/OL]. (2020 – 10 – 20) [2025 – 07 – 08]. https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_EN.html.

[6] 陈吉栋. 以风险为基础的人工智能治理 [J]. 法治研究，2023 (5): 52 – 60.

[7] 胡巧莉. 人工智能服务提供者侵权责任要件的类型构造——以风险区分为视角 [J]. 比较法研究，2024 (6): 57 – 71.

[8] 张凌寒. 论数据信息损害的承认与救济 [J]. 中国法学，2024 (3): 62 – 82.

[9] European Commission. Proposal for a directive on adapting non-contractual civil liability rules to artificial intelligence: COM (2022) 496 final [EB/OL]. (2022 – 09 – 28) [2025 – 07 – 08]. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0496.

[10] LA DIEGA G N, BEZERRA L C T. Can there be responsible AI without AI liability? Incentivizing generative AI safety through ex-post tort liability under the EU AI liability directive [J]. International Journal of Law and Information Technology, 2024, 32 (1): eaae021.

[11] 戴昕. 无过错责任与人工智能发展——基于法律经济分析的一个观点 [J]. 华东政法大学学报，2024，27 (5): 38 – 55.

[12] 崔国斌. 著作权法：案例与原理 [M]. 北京：北京大学出版社，2014.

[13] 程啸. 人格权研究 [M]. 北京：中国人民大学出版社，2022.

[14] CHUN J, SCHROEDER DE WITT C, ELKINS K. Comparative global AI regulation: policy perspectives from the EU, China, and the US [EB/OL]. (2024 – 10 – xx) [2025 – 07 – 08]. https://arxiv.org/abs/2410.21279.

[15] European Commission. Commission work programme 2025: Moving forward together: a bolder, simpler, faster union: COM (2025) 45 final [EB/OL]. (2025 – 01 – 22) [2025 – 07 – 08]. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52025DC0045.

[16] 张凌寒. 中国需要一部怎样的《人工智能法》？——中国人工智能立法的基本逻辑与制度架构 [J]. 法律科学（西北政法大学学报），2024，42 (3): 3 – 17.

[17] SHAVELL S. Liability for harm versus regulation of safety [J]. The Journal of Legal Studies, 1984, 13 (2): 357 – 374.

[18] 朱虎. 规制法与侵权法 [M]. 北京：中国人民大学出版社，2018.

[19] European Union. Artificial Intelligence Act：Regulation（EU）2024/1689［EB/OL］.（2024－07－12）［2025－07－08］. https：//eur-lex. europa. eu/legal-content/EN/TXT/？uri = CELEX%3A32024R1689.
[20] CLAASSEN R. Externalities as a basis for regulation：a philosophical view［J］. Journal of Institutional Economics，2016，12（3）：541－563.
[21] ZHANG A H. High wire：How China regulates big tech and governs its economy［M］. Oxford：Oxford University Press，2024：232－238.
[22] 程啸. 论侵害个人信息的民事责任［J］. 暨南学报（哲学社会科学版），2020，42（2）：39－47.
[23] 阿瑟·塞西尔·庇古. 福利经济学［M］. 朱泱，张胜纪，吴良健，译. 北京：商务印书馆，2017.
[24] 凌斌. 法律救济的规则选择：财产规则、责任规则与卡梅框架的法律经济学重构［J］. 中国法学，2012（6）：5－25.
[25] 崔国斌. 知识产权法官造法批判［J］. 中国法学，2006（1）：144－164.
[26] 崔国斌. 网络服务商共同侵权制度之重塑［J］. 法学研究，2013，35（4）：138－159.
[27] United States Congress. S. 146-Take It Down Act：119th Congress（2025—2026）［EB/OL］.（2025－05－19）［2025－07－08］. https：//www. congress. gov/bill/119th-congress/senate-bill/146.

（收稿日期：2025－07－08）

作者简介：

徐榕岭（2000－），女，硕士，助理研究员，主要研究方向：知识产权法、网络法。

（上接第74页）

[3] 刘继峰，张佳红. 平台封禁行为的竞争法分析［J］. 当代经济科学，2023，45（1）：17－28.
[4] 魏远山. 我国反不正当竞争法商业数据专条的制度构建——兼评《反不正当竞争法（修订草案征求意见稿）》第18条［J］. 环球法律评论，2023，45（6）：80－96.
[5] 王永强，邓淑元. 企业数据保护反不正当竞争法路径的展开［J］. 中南民族大学学报（人文社会科学版），2024，44（8）：134－141，186－187.
[6] 丁晓东. 论数据垄断：大数据视野下反垄断的法理思考［J］. 东方法学，2021（3）：108－123.
[7] 冯博，张家琛. 数字平台领域《反垄断法》与《反不正当竞争法》的经济逻辑和司法衔接［J］. 法治研究，2023（3）：83－94.
[8] 李鑫. 平台数据型自我优待的反垄断法分析［J］. 广东财经大学学报，2023，38（1）：101－111.
[9] 杨东，李子硕. 论反不正当竞争法的重构：以数据要素规制体系为核心［J］. 法律适用，2022（11）：15－25.
[10] 王永强. 包容审慎监管视角下平台经济竞争失序的法治应对［J］. 法学评论，2024，42（2）：77－87.
[11] 黄锫. 论我国人工智能领域包容审慎监管的法治维度［J］. 财经法学，2025（2）：94－109.
[12] 王永强，李佐. 商业数据法际协同治理的规范路径及其展开［J］. 湖北经济学院学报，2025，23（2）：80－90.
[13] 宋亚辉. 论反不正当竞争法的一般分析框架［J］. 中外法学，2023，35（4）：963－982.
[14] 曹汇. 涉数据不正当竞争行为规制研究［J］. 大连理工大学学报（社会科学版），2024，45（3）：87－96.
[15] 周围，黄唯一. 反垄断法与隐私保护法的矛盾与纾解［J］. 华中科技大学学报（社会科学版），2023，37（1）：97－107.
[16] 王延川. 数据法人：超级平台数据垄断的治理路径［J］. 国家检察官学院学报，2022，30（6）：145－159.
[17] 孔祥俊. 商业数据保护的实践反思与立法展望——基于数据信息财产属性的保护路径构想［J］. 比较法研究，2024（3）：72－92.
[18] 许光耀. 大数据杀熟行为的反垄断法调整方法［J］. 政治与法律，2024（4）：17－29.

（收稿日期：2025－06－20）

作者简介：

李佐（2000－），男，硕士研究生，主要研究方向：经济法学、数字法学。

版权声明

凡《网络安全与数据治理》录用的文章，如作者没有关于汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权等版权的特殊声明，即视作该文章署名作者同意将该文章的汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权授予本刊，本刊有权授权本刊合作数据库、合作媒体等合作伙伴使用。同时，本刊支付的稿酬已包含上述使用的费用，特此声明。

《网络安全与数据治理》编辑部