财经法学 No. 5，2025
pp. 115 - 130

从限制到调控：人工智能背景下未成年人保护模式的转型与重塑

刘晓春*

内容提要：未成年人使用人工智能应用的行为具有高度交互性、私密性、工具性和枢纽性等特点，对未成年人保护提出了新挑战。我国未成年人模式已经取得的实践成果，在实施效果上依然面临质疑。我国未成年人模式的运行机理包括身份识别、权限设置和家长赋能三个方面，具有明显的限制型特征，面临输入、处理、输出三个层次的信息困境，导致保护要求难以落地。人工智能背景下，需要将未成年人模式进行改造，转向调控型保护模式，即基于人工智能技术带来的信息处理能力提高和成本降低，破除信息困境，从身份识别、风险防范、发展促进、家长赋能四个层次展开系统性重塑，并建构相应的制度保障，实现个性化、精准化、赋能化的未成年人保护。

关键词：未成年人网络保护　未成年人模式　人工智能　调控型保护

一、问题的提出

随着人工智能时代的快速到来，未成年人使用人工智能应用和工具的规模和比例快速上升，未成年人保护制度的建构和实施持续面临挑战，并受到国内外公众和政策制定者的高度关注。我国的未成年人网络保护制度在基础立法上已经完成体系化建构，以《中华人民共和国未成年人保护法》（2024 年修正，以下简称《未成年人保护法》）网络专章和《未成年人网络保护条例》相继通过并实施作为标志性事件，相应工作机制的完善和落地已有比较完整的法律规范，也为国际上其他国家提供了重要范例。

未成年人保护模式是我国颇具特色的保护机制，在实施过程中不断演化和迭代，目前覆盖面

* 刘晓春，中国社会科学院大学法学院副教授、互联网法治研究中心主任。

包括未成年人使用网络的几乎所有领域，功能也从最初的“游戏防沉迷”定位不断扩充，目前已涵盖了时间管理、内容管理、权限管理、消费管理等多元功能。〔1〕2024 年 11 月由网信部门发布的《移动互联网未成年人模式建设指南》，是监管部门近期对于未成年人模式建设的重要举措，通过细化具体建设要求，力求推动软硬件多方联动、便捷使用、分龄设置、家长赋能等制度目标的有效实现。〔2〕

与此同时，未成年人模式对于未成年人群体的限制型保护思路，也面临落地的难点。不过，这一思路在最近一段时期恰恰在国际上被积极推行，在澳大利亚、欧洲、美国加州都出现了限制特定年龄未成年人使用社交媒体的立法动态。〔3〕这种模式可以通过限制甚至禁止未成年人使用特定功能来控制和消除风险，保证未成年人不受网络空间潜在风险的侵害。但是，无论在网络时代还是人工智能时代，必须承认的一个现实是，已经不太可能把未成年人这群网络“原住民”与网络和人工智能技术有效区隔开来，未成年人逃离以未成年人模式为代表的“限制型”模式的行动，从来没有停止过。其实际的适用效果就是，大量未成年人以成年人身份在网络上活动，平台缺乏有效的识别机制，未成年人模式在很多时候被质疑为“形同虚设”。

正在快速发展的人工智能应用是否应当继续沿用未成年人模式，还是应当另辟蹊径，成为当下思考和探讨的重要问题。有必要重新考虑现有未成年人模式的不足、困境和成因，重新配置政府、平台、家长等各方主体的角色和义务，并充分结合人工智能技术在治理和保护方面带来的可能性，推动人工智能时代未成年人保护模式的转型与重塑。

二、人工智能应用的未成年人使用特点与保护难题

在以大语言模型为代表的生成式人工智能技术不断演进过程中，涌现了大量具有鲜明人工智能特点的网络应用，使用人工智能应用的未成年人用户数量也呈现快速增长的态势，未成年人使用人工智能应用涉及的风险和保护问题引发公众高度关注。由于人工智能应用的发展方兴未艾，就未成年人使用这些应用的特点、风险和治理也尚未达成较为成熟的共识，通过何种具体保护模式来实现对未成年人的保护，是否需要针对人工智能应用设置未成年人模式，原有的互联网未成年人模式是否能够实现有效的保护并促进未成年人发展，都是人工智能技术快速发展背景下需要进行研究和回答的问题。

（一）人工智能应用服务中未成年人使用行为的特点

近年来取得明显进步的人工智能技术，具有日益增强的对于自然语言进行理解、生成、推理、决策等的能力，在为各行各业的智能化转型进行技术赋能的同时，也支持推出面向终端用户的各种人工智能具体应用，通过提供聊天互动、生成合成、推理演算、辅助决策等功能，展现了

〔1〕 参见林维、吴贻森：《网络保护未成年人模式：立法跃升、理念优化与困境突破》，载《吉林大学社会科学学报》2022 年第 5 期。

〔2〕 参见《移动互联网未成年人模式建设指南》，载 https://www.gov.cn/lianbo/bumen/202411/content_6987450.htm，最后访问时间：2024 年 12 月 15 日。

〔3〕 参见央视新闻：《关于未成年人合理上网，多国这样做》，载 https://content-static.cctvnews.cctv.com/snow-book/index.html?item_id=6110807010627526315，最后访问时间：2025 年 1 月 3 日。

广阔的商业应用前景，对包括大量未成年人在内的用户产生非常明显的吸引力，用户数量快速攀升。〔4〕在人工智能应用快速发展之前，网络游戏、短视频、网络直播等内容服务以及网络社交服务属于未成年人使用频率最高、程度最深的服务类型，也是《未成年人保护法》和《未成年人网络保护条例》等立法重点关注的领域。〔5〕与这些服务类型相比，人工智能应用服务中未成年人的使用行为具有比较明显的特点。

首先，人工智能应用与用户之间通常存在比较密集的个性化交互。这在目前作为人工智能应用主流类型之一的聊天类应用中表现得最为明显，通常被认为具有“陪伴型”的属性。在交互过程中，用户一方面更有可能向人工智能应用输入高度个性化的信息，另一方面也可能通过较为密切深入的交流，受到人工智能应用输出信息、内容、观点、建议的更深刻的影响。对于这种影响，如果说成年人基于其较为成熟的认知能力通常还是可以作出独立、理性判断的话，那么对于心理和认知能力尚不成熟的未成年人来说，就更容易受其引导和干扰。陪伴型人工智能对未成年人的心理和价值观等层面的影响，目前虽然较难断言已经导致较高风险，但零星出现的极端事件，已经引起公众的警惕和担忧。〔6〕相比之下，网络游戏、短视频、直播等业态，都是在内容生产的基础上面向不特定人或者多数人进行传播，尽管在算法推送机制上具有个性化推荐的属性，但是无论从信息输入还是内容传播来看，都无法达到聊天和陪伴型人工智能应用模式下“一对一”的高度个性化程度。

其次，人工智能应用中的用户使用行为具有私密性。这一特征同样在聊天、陪伴型的应用中最为明显，而同样适用代表未来发展趋势的人工智能私人助理等方面。在游戏、短视频、直播等主要以公开形式传播的内容服务中，可以通过事前、事中、事后的审核、干预、提示、调控等方式来实现内容治理。〔7〕而人工智能应用在私人空间提供的大多是即时生成的内容服务，使得对其的审核和调整很难及时有效作出，而只能主要通过事前的技术设计和测试规范来实现。服务提供私密性这一特征与网络社交场景有所类似，但后者的未成年人使用风险主要来自与真人的互动，与人工智能应用呈现不同的特征。

再次，人工智能应用具有工具性的特征。在网络游戏、短视频、直播等领域，未成年人用户主要是内容服务的被动接受者，仅通过内容搜索、选择、评论等行为掌握一定的主动性。在使用人工智能应用的生成合成、推理决策等功能时，用户往往能够通过这些智能工具形成新的内容，从生成文字、图片、音乐、视频到编程、研究、设计新的智能体，人工智能应用正在迅速为用户提供众多领域的增强功能。对于未成年人而言，这些工具是通向人工智能时代的重要桥梁，大大拓展了未成年人的技能和想象空间，使得他们正在越来越从内容的被动接受者转化为内容的主动

〔4〕 参见中国信通院：《人工智能发展报告 2024 年》，载 http://www.caict.ac.cn/kxyj/qwfb/bps/202412/P020241210548865982463.pdf，最后访问时间：2025 年 1 月 3 日。

〔5〕 参见卢家银、邹琴：《多元共治下的以网管网：中国特色未成年人网络保护的实践模式》，载《少年儿童研究》2024 年第 2 期。

〔6〕 参见赵丽、马子煜：《与孩子聊天的 AI 人设是“出轨对象”？记者调查 AI 剧情聊天软件乱象》，载《法治日报》2024 年 11 月 23 日，第 5 版；法治网：《一款 APP 被约谈，曾和未成年人聊色情！》，载 https://mp.weixin.qq.com/s/gcRBz4_bs9d4IPCkmuzKDA，最后访问时间：2025 年 6 月 22 日。

〔7〕 参见支振锋、刘佳琨：《互联网信息内容治理的中国方案》，载《江西社会科学》2023 年第 11 期。

生产者和传播者。〔8〕而人工智能的工具性特征，既能为未成年人带来积极效果，也可能带来风险和弊端，成为一柄双刃剑。未成年人通过人工智能工具拓展数字能力和素养的同时，也可以将其用于不良甚至有害内容的生产制作之中。〔9〕

最后，人工智能应用为用户提供枢纽性的服务，通过搜索、推荐、链接等方式，将不同种类平台的服务整合在一起，逐步成为用户访问网络的入口。这是人工智能助手发展的未来方向，特别是在移动终端的端侧人工智能领域迅速发展，并在目前的各款综合类人工智能应用中体现得较为明显。对于未成年人来说，通过与人工智能应用进行个性化、私密性的互动，可以通过点击被推荐的链接等方式获取不同平台公开传播的内容，从而通过人工智能的枢纽性功能获得跨平台的内容推荐组合。与同样具有枢纽和跨平台聚合性的搜索引擎不同，人工智能应用具有更强的主动推荐性，并且有可能与生成内容结合，对于用户需求的反应更具针对性，并以中心化的方式呈现，同时在风险和不良影响层面也可能具有跨平台整合的影响。

未成年人使用人工智能应用体现出来的交互性、私密性、工具性和枢纽性等特征，使得人工智能背景下的未成年人保护面临新的问题和挑战。与此同时，未成年人网络保护长期以来未能解决的难题，也会在人工智能背景下延续，需要结合人工智能应用的特点给出整体的解决方案。

（二）保护难题：从网络保护到人工智能应用

我国的未成年人网络保护制度经过多年的实践探索，目前已经建立起比较完整的立法体系，以《未成年人保护法》和《未成年人网络保护条例》为基础，在总结和提炼实践经验的前提下取得了重要的立法成就。〔10〕如何将立法的原则和目标通过具体的工作机制落地实施，成为下一个阶段的重点任务。〔11〕面向人工智能时代的到来，未成年人网络保护面临的挑战既包括原有制度的实施难题，也包括人工智能应用带来的新型问题。随着技术的变迁，未成年人网络保护越来越超越家庭监护的内部范畴，成为需要政府、平台、学校和社会各界共同努力的事业。而针对未成年人群体的区分保护和特殊保护何以可能，在人工智能背景下是否需要调整，未成年人模式作为我国未成年人网络保护的重要制度成就是否应当直接适用于人工智能应用领域，其效果如何评估，都是需要直面和回应的难题。〔12〕

1. 家庭监护的困境与未成年人保护的公共化

未成年人的照看和养育，传统来看是比较典型的家庭事务，主要通过监护制度、婚姻家庭相关制度等维护私法领域的利益秩序，除非出现监护人缺位、家庭暴力等特殊情况，一般而言政府机关和社会组织不会直接介入家庭内部的亲权监护关系，从而确立了家庭保护在未成年人保护中的主导地位。

---

〔8〕 参见林维、刘晓春：《构建面向未成年人保护的网络生态协同治理体系》，载《中国党政干部论坛》2023 年第 12 期。

〔9〕 参见赵丽、马子煜：《与孩子聊天的 AI 人设是“出轨对象”？记者调查 AI 剧情聊天软件乱象》，载《法治日报》2024 年 11 月 23 日第 5 版。

〔10〕 参见林维、刘晓春：《中国未成年人网络保护发展现状与展望》，载林维主编：《未成年人网络保护发展报告（2021）》，中国社会科学出版社 2022 年版，第 8－10 页。

〔11〕 党的二十届三中全会审议通过的《中共中央关于进一步全面深化改革、推进中国式现代化的决定》对健全网络综合治理体系作出了明确要求，并特别提到要“健全未成年人网络保护工作体系”。

〔12〕 参见林维：《网络保护未成年人模式的强化与推行》，载《少年儿童研究》2021 年第 8 期。

但是，未成年人网络使用过程中产生的诸多风险与挑战，伴随着公众关注和对策讨论，使得未成年人网络保护问题逐步从家庭监护领域为主的问题，演化为全社会关注的公共事务。家庭监护在未成年人网络保护领域存在的困境，以及基于此产生的公共干预的需求，可以从三个方面来理解。

首先，未成年人使用网络不再受限于物理空间或现实社交互动，数字化行为使得家庭不再构成与外界区隔的独立物理空间，因此传统线下通过家长陪伴或代理而对外交往的模式，很容易被绕过从而难以构成未成年人与外界接触的物理“屏障”。其次，在数字化生活中，家长对未成年人的控制方式和能力十分有限，甚至出现了部分未成年人网络运用能力超过家长、能力倒挂的情况，家长的监护和管理行为与未成年人的自主性和权益诉求出现明显冲突，比如家长对未成年人手机使用的监控需求与未成年人对隐私保护的强烈主张之间的矛盾。〔13〕最后，对于国家和社会而言，未成年人网络保护问题成为公共舆论持续关注的焦点问题，与传统线下的未成年人侵害较多呈现为孤立事件不同，网络上的未成年人保护和发展问题往往具有普遍性。因此，未成年人网络保护作为公共事务在网络空间治理中具有较高优先级，应由政府、社会共同投入实质资源的共识也逐步形成。

未成年人网络保护的公共化，使得政府承担起较高的公众期待，一定意义上也构成了国家亲权的行使依据。〔14〕但是，要求政府全面负担起对于具体未成年人在网络使用中的保护职能，无论从资源投入成本和实际可操作性上，都存在不切实际之处。因此，基于我国网络空间治理的基本体系和框架，政府将实质性的保护任务通过立法等方式交予网络平台，要求其将未成年人保护作为法定义务加以履行，从而构成了一种政府与平台之间的发包结构。〔15〕由此，在重申家庭、学校、政府、社会在未成年人保护中各自角色的同时，网络平台实际上在其中承担起了至关重要的角色。《未成年人保护法》修订中加入“网络保护”专章，以及《未成年人网络保护条例》的体系构建，都充分反映了这一理念和思路。〔16〕但是在制度落地过程中，对未成年人的有效识别、特殊保护、发展引导这些机制都存在尚未克服的难题，这些问题集中体现在未成年人保护模式这一最为重要的平台举措之上。

2. 未成年人模式的保护实效尚存争议

未成年人模式作为具有特色的未成年人保护机制，在实践中的推行取得了重要成果，但也存在需要改进和完善之处。目前，这一模式的基本理念是“隔离＋限制”，即通过识别出未成年人并对其进行群体“隔离”，通过对未成年人用户使用权限的限制甚至取消，来降低风险，通过减少乃至消除接触和使用网络的可能性，来防止对未成年人身心健康的可能侵害。这一做法的确可能起到保护效果，但是也带来迄今为止未能克服的适用困境。

首先，“隔离”和“限制”的基本方式容易引起未成年人的抵触心理，导致他们有非常强烈

〔13〕 参见陆杰华、谷俞辰：《父母网络行为干预对未成年人网络重度使用倾向的影响探析》，载《华中科技大学学报（社会科学版）》2023年第5期。

〔14〕 参见何波：《论中国未成年人网络保护法律制度体系的改进》，载《法律科学（西北政法大学学报）》2024年第2期。

〔15〕 参见胡凌：《平台发包制：当代中国平台治理的内在逻辑》，载《文化纵横》2023年第4期。

〔16〕 参见林维、刘晓春：《中国未成年人网络保护发展现状与展望》，载林维主编：《未成年人网络保护发展报告（2021）》，中国社会科学出版社2022年版，第8－10页。

的冲动“逃离”未成年人模式。特别是在模式内的时间、内容和功能受到严格限制的情况下，未成年人或者是弃用服务，或者是转而注册新账号并想方设法避免家长将其“关入”未成年人模式，形成“猫鼠游戏”以及经常被提到的未成年人模式“形同虚设”。其次，家长出于各种原因，对于未成年人模式的了解、设置、管理都尚未形成较为普遍的积极性、主动性和专业性。在缺乏家长和未成年人主动启动行为的情况下，平台亦无激励主动识别并启用未成年人模式。第三，在大量未成年人游离于未成年人模式之外的情况下，平台除了基于合规需求而完成未成年人模式建设的最低标准之外，并无强烈的激励对于未成年人模式的内容池和功能权限进行更加进阶的设计，进而导致对于未成年人的吸引力无法真正提高。由此，在很多情况下，未成年人模式成为网络平台满足合规形式要求的“花瓶式”设计，大多数未成年人并未被有效识别出来，反而以成年人用户身份使用平台。〔17〕

在人工智能应用场景下，目前比较常见的类型包括陪伴型、生成型、工具型等服务。由于目前国内外人工智能应用大都尚未建立未成年人特殊保护机制，对于人工智能应用是否应当采用未成年人模式的讨论也已经提上议事日程。在人工智能时代，是继续沿用以隔离和限制为主要思路的未成年人模式，还是充分运用人工智能技术提供的可能性，对未成年人模式进行理念更新和模式重塑，并修改和重构支撑性的法律规则，是无法回避的时代问题。

三、限制型未成年人保护模式的运行机理与信息困境

在未成年人保护实践中逐步发展演化的未成年人保护模式，可以认为是我国未成年人网络保护的一个特色制度，通过为未成年人提供特殊的服务内容和模式来实现特殊保护的目标，并且试图通过特定的设置功能为家庭监护赋能，把家长也纳入“政府—平台”的治理结构之中，形成“政府—平台—家长”的三元治理模式。〔18〕但是未成年人模式在实践中依然存在落实难题，特别是，现有的未成年人模式主要是通过限制未成年人使用服务的方式来实现保护，构成“限制型”保护模式，这一模式在保护实效方面的争议主要来源于运行过程中的信息不足的困境，而以限制使用为特点的模式在运行过程中，进一步加重了信息困境。

（一）运行机理

未成年人模式的运行机理可以总结为身份识别、权限设置、家长赋能三个层次，其基本思路是将未成年人作为特殊群体进行类型化的识别，进而通过限制其相关权限控制风险、预防侵害，并通过技术手段辅助家长实现对于未成年人更加有效的指导和保护。

1. 身份识别

针对未成年人提供特殊保护的前提是能够把未成年人用户有效区分出来。我国目前实行网络实名制，要求用户在进行信息发布、即时通信时需要提供真实身份信息，但是除了支付、游戏等少数领域存在身份证实名等强实名要求之外，大多领域主要是采用手机号等较弱的身份认证方

〔17〕 参见林维、刘晓春：《构建面向未成年人保护的网络生态协同治理体系》，载《中国党政干部论坛》2023 年第 12 期。

〔18〕 参见卢家银、邹琴：《多元共治下的以网管网：中国特色未成年人网络保护的实践模式》，载《少年儿童研究》2024 年第 2 期。

式，并且实行“后台实名、前台自愿”的规则。〔19〕对于网络平台而言，未成年人用户有可能基于两方面原因而难以识别。一是手机号为基础的身份验证信息不足以体现未成年人实际用户的特征，二是现实中大量未成年人并非以自己真实身份信息进行平台账号注册，例如最为常见的是使用注册在家长名下的手机号注册平台账号并使用，也可能存在同一平台账号家长和未成年人同时使用的情况。

目前，未成年人模式运行中的未成年人识别主要通过三种方式来实现，一是未成年人主动披露真实身份，例如通过提供身份证号进行强实名认证，这在网络游戏中存在明确要求，而在短视频、直播等其他平台则较为少见，后者通常是通过手机号实现弱实名认证。〔20〕二是家长通过对未成年人账号的控制和干预，而将其设置为未成年人模式，从而通过家长主动披露实现未成年人身份识别的效果。三是平台主动对未披露身份的未成年人用户根据其行为习惯和特征进行推测和判断，对疑似未成年人账号进行不同程度的识别和权限设置，并通过要求用户进行特殊验证方式进行确认，例如进行强化的身份信息验证等。〔21〕

2. 权限设置

未成年人模式的技术本质就是针对用户账号实施权限设置，即对于识别出来并加入未成年人模式的用户设置特殊权限，通常是相对于普通用户而言施加限制性乃至禁止性措施，包括时间、内容、功能等方面的限制和管理。通过这些限制措施，未成年人模式被认为可以消除未成年人身心健康受侵害的风险，使未成年人能够在安全的环境中使用网络。

一是时间限制，这类措施源于未成年人模式早期被着重强调的防沉迷功能，包括时长和时段两个方面的限制，这在游戏类应用中最为明确和严格，即仅在周五、周六、周日和法定节假日每日 20—21 点向未成年人提供 1 小时网络游戏服务。〔22〕时间设置特别是时长的限制也逐步应用到音视频、社交媒体等其他领域的应用。针对不同应用的时长限制无法打通和共享的问题，最新的政府指引文件要求在移动终端中将所有应用的时间限制统一进行设置，以防止未成年人在不同应用间切换导致整体使用时间依然过久的问题。〔23〕

二是内容限制，即未成年人模式中的内容需要经过区别于普通模式标准的特殊筛选，建立专属内容池，防止有害和不良内容对未成年人的可能侵害。〔24〕目前业界通过“黑名单”和“白名单”两种方式来实现内容筛选。〔25〕这两种模式下，“黑名单”模式的内容数量和丰富程度通常远

〔19〕 参见任秀、王咏梅：《全网实名制后的困境与对策研究》，载《湖北社会科学》2020 年第 4 期。

〔20〕 参见任秀、王咏梅：《全网实名制后的困境与对策研究》，载《湖北社会科学》2020 年第 4 期。

〔21〕 参见互联网法学 SKD：《“网络游戏未成年人保护：指标测评与机制完善”研讨会顺利举行》，载 https://mp.weixin.qq.com/s/91uShVAoJT4QXFBJF1GMhw，最后访问时间：2024 年 12 月 20 日。

〔22〕 参见《国家新闻出版署关于进一步严格管理切实防止未成年人沉迷网络游戏的通知》第 1 条。因其规定严格，被媒体称为“史上最严防沉迷”规定。

〔23〕 参见《移动互联网未成年人模式建设指南》，载 https://www.gov.cn/lianbo/bumen/202411/content_6987450.htm，最后访问时间：2024 年 12 月 15 日。

〔24〕 参见李红勃、张玉芳：《未成年人保护视野下网络信息内容规范管理的制度设置与优化建议》，载《中国青年社会科学》2024 年第 1 期。

〔25〕 参见彭桂兵、胡钊涵：《发现与处置：短视频平台在未成年人保护中对内容传播的管理义务》，载《西南民族大学学报（人文社会科学版）》2024 年第 10 期。

超“白名单”模式，由此对未成年人用户的使用体验和吸引力也存在明显差异。此外，有些内容直接根据其展示形式就可能会直接被判定为存在风险，不在未成年人模式内展现，如直播、微短剧等经常在舆论中出现争议的形式，就不再对实质内容进行评价，实行所谓“一刀切”的屏蔽。实践中，大部分平台采用的是更加安全的“白名单”模式，而且经常出现内容种类单一、数量有限、对未成年人缺乏基本吸引力的情况，其诉求是在形式上满足合规要求。因此，此类未成年人模式中实际使用情况存疑，也经常被媒体质疑为“形同虚设”。〔26〕

三是功能限制，即未成年人模式会对用户的部分功能进行限制，以防范可能带来的风险。除了《未成年人保护法》和《未成年人网络保护条例》等文件明确提到的消费限制、算法推送、广告营销等限制要求之外，〔27〕网络平台通常还会对未成年人模式下的信息发布、社交互动等功能进行限制甚至取消，以确保未成年人不会暴露在这些行为带来的风险当中。但是，跟内容限制的情况一样，网络平台基础和核心功能的“阉割”式限制，也使得未成年人模式里面的使用体验与用户期待相去甚远，无法真正留住未成年人用户，一旦被“关进”未成年人模式，也就相当于要求未成年人弃用平台服务。

3. 家长赋能

未成年人模式发挥作用，除了需要平台提供相应的身份识别和权限设置功能外，在很大程度上也需要家长的主动使用作为基础。身份识别方面，未成年人主动披露和平台识别的比例总体上较低，大多数情况下，需要家长主动开启未成年人模式。在权限设置方面，家长也应在平台提供的默认设置基础之上进行个性化的选择，从而实现最适合孩子的权限配置。

从另一个视角出发，未成年人模式也是在为家长更好地引导和保护未成年人用网行为提供技术赋能。家长监护能力的不足，具体而言主要体现在信息获取能力和数字保护技能上。信息获取能力既包括家长对未成年人网络使用行为信息的了解，也包括对于网络可能存在侵害的相关信息的获取。数字保护技能指的是家长在数字环境下如何实现对于未成年人的有效引导和保护的能力。理想状态下，未成年人模式通过配置家长“守护”权限，可以使家长在简单地将未成年人“关进”该模式之外，还能了解未成年人网络使用情况，并进行相应的内容、功能权限设置，例如将特定内容类别自行选择加入“白名单”或移出“黑名单”，根据未成年人的年龄和认知水平调整消费管理的限额，以及对于社交等功能进行定向开放和调试等。

（二）“政府—平台—家长”治理框架下的各方分工协同

在未成年人模式运行过程中，通过厘清政府、平台、家长三方在未成年人网络保护中的角色定位和分工协同，可以更加清楚地观察以限制为特点的未成年人模式在实践中运行的状况及其不足。

首先是政府。政府在未成年人模式制度建设中的角色可以从三个层面来观察。一是机制建设

〔26〕 参见雷雳、王兴超：《网络平台青少年模式缘何形同虚设》，载《人民论坛》2020 年第 28 期。

〔27〕 关于消费功能的限制，《未成年人网络保护条例》第 44 条规定：“网络游戏、网络直播、网络音视频、网络社交等网络服务提供者应当采取措施，合理限制不同年龄阶段未成年人在使用其服务中的单次消费数额和单日累计消费数额，不得向未成年人提供与其民事行为能力不符的付费服务。”关于算法推送商业营销的限制，该条例第 24 条第 3 款规定：“网络产品和服务提供者不得通过自动化决策方式向未成年人进行商业营销。”

中基础规则的提供。在未成年人保护上，政府未必有能力和资源事必躬亲地介入每一个个案，但是在基础规则制定上存在信息和能力上的优势。政府不仅要求设立未成年人模式，要求平台积极引导和宣传用户使用未成年人模式，还对未成年人模式的具体功能和细化规则提出要求和指引。二是机制运行中的动态指导，尤其是对于不断涌现的新问题新风险，政府有能力进行整体态势研判和专项指导，面向不同平台给出动态的高风险场景指引和专项治理要求，并在一定程度上统一治理标准，如“清朗”系列行动。三是协调确定利益冲突下的价值排序以及不同主体间的权责分配。

其次是平台。在网络空间治理中，平台通常被认为是拥有较强的信息和技术能力，并能够以相对可控的成本展开各项治理的主体。〔28〕就未成年人模式的运行看，对于平台而言，模式设计、账号处置、家长赋能这些确定性较强的主要为技术层面的任务，属于平台以可控成本即可履行的义务。但是就身份识别而言，如果是未成年人或其家长主动加入未成年人模式，则平台有能力以较低成本进行保护，但如果是要求平台自行对前台匿名的未成年人用户进行精准识别，则往往超出了平台的现有能力范围。〔29〕平台如果需要落实主动进行身份识别的功能，需要承担三方面的成本。一是开发专门技术模型或采取技术措施进行用户画像分析和判断的成本，例如目前有些游戏平台采取的在技术模型分析判断基础上进行刷脸认证的方式；二是大规模识别措施导致误伤并进而损害用户体验的成本；三是强化的用户数据分析过程可能还面临着个人信息保护合规方面的隐忧，例如刷脸认证极易引发公众对于数据泄露风险的担忧，因而目前即使在游戏领域也并未进行广泛应用。从激励角度，网络平台的未成年人网络保护激励主要来自法律合规要求和基于社会责任的声誉维护。因此，对于平台而言，一方面有能力也有意愿设置未成年人模式并配备基本功能，以满足法律合规要求，另一方面，又倾向于在模式里被动接收未成年人用户而非主动识别，对于内容池的建设也以确保安全为主要诉求，倾向于控制内容数量、种类和尽量限制存在风险的功能选择。这也是形成大量“形同虚设”未成年人模式的背后动因。〔30〕

最后是家长。家长通常被认为是未成年人保护的第一责任人，但是这在网络时代却面临众多难题。从成本看，家长作为离未成年人最近也是最容易了解未成年人情况的主体，在个性化保护方面相对于政府与平台具有明显的优势。但是在实践中，家长在网络时代的监护能力的确面临挑战。家长在难以轻易找到有效监护手段的情况下，倾向于对平台和政府提出更多期望，要求其承担更多责任。例如在未成年人大额消费情况下，家长也经常主张平台负有全责从而应当全额退款，而不认为自己存在疏忽过错。未成年人模式中“家长赋能”的环节，即意在通过增加家长的能力降低其成本来提升保护效果。此外，在留守儿童等情形下，家长的能力、成本等又呈现不同的特点，在制度设计和责任配置上也需要考虑到此类特殊情况。

（三）现有模式下的信息困境和制度成因

以限制使用为主要特征的未成年人保护模式，就其运行机理来说，试图通过识别和隔离来对未成年人进行区分，进而通过设置限制型的权限来控制风险，并通过向家长提供个性化设置的技

〔28〕 参见邵聪：《未成年人网络保护中的企业合规实践与制度完善》，载《中国青年社会科学》2024年第1期。

〔29〕 参见林维、刘晓春：《构建面向未成年人保护的网络生态协同治理体系》，载《中国党政干部论坛》2023年第12期。

〔30〕 参见雷雳、王兴超：《网络平台青少年模式缘何形同虚设》，载《人民论坛》2020年第28期。

术辅助手段来平衡限制与发展之间的关系。这一看似妥当的逻辑却在实践运行中面临被未成年人主动规避和逃离的质疑。究其根本，通过限制型理念来对未成年人进行保护的模式无法获得未成年人的认同，也难言是符合最有利于未成年人原则的制度设计。〔31〕之所以在实践中形成了限制型的理念和路径，恰恰是因为整个保护机制建立在信息不足的困境之上，可以说有其明显的制度成因。

未成年人保护的制度目标主要在于控制风险和促进发展，这与很多以控制风险为主要导向的治理制度有所不同。无论是控制风险还是促进发展，都需要能够基于充分而准确的信息进行判断和决策。涉及风险控制的治理路径大体可以分为两种模式，一种是通过控制和减少信息处理和供应来降低风险，例如事前的内容审查、广告审查、个人信息最小化原则等，都属于此类。第二种是通过提供和处理更加充分的信息来降低风险，例如信息披露、广告标识、信用评价等制度。平台治理过程中逐步形成的“调控型”治理模式，即属于后者，例如，在直播带货的治理中，“调控型”治理模式基于实时检测技术、数据处理和算法分析，运用信用工具、信息工具以及平台增信工具等基础设施，激励行为主体生产更多数量的可信信息，由此建立新型信任机制并控制风险。〔32〕

在未成年人网络保护中，限制型的未成年人保护模式就面临着三个层次的信息困境，分别涉及信息收集、信息处理和信息输出三个阶段。

就信息收集而言，对未成年人进行有效保护，至少需要获取其身份和行为两类信息。身份识别信息是有效区分未成年人并进行特殊保护的前提，但是在现有的未成年人模式下，这类信息的获取无论是在数量还是在质量上都十分欠缺。如前所述，目前未成年人身份识别主要基于未成年人主动披露、家长披露以及平台根据使用行为进行识别三种途径。在这个过程中，由于未成年人模式的限制型特征，未成年人整体上缺乏主动披露的意愿，往往冒充成年人身份提供信息，这就造成真实身份信息的严重缺失。在家长披露的情况下，也会出现未成年人或者对未成年人模式弃而不用，或者想方设法逃离未成年人模式的情况，依然会造成信息困境。平台主动识别的行为涉及对未成年人个人信息的收集和处理，面临合法性基础的疑问，因此平台也并无积极通过此种途径获取未成年人特征信息的激励。

就信息处理而言，包括用于识别和匹配两种目的的处理行为。用于识别目的，指的是处理未识别用户的信息用于识别未成年人；用于匹配目的，指的是对于已识别的未成年人用户的特征进行分析从而用于推荐适合的内容。这两种情况都会涉及形成并运用未成年人用户画像的问题。实际上，用户画像在移动互联网时代具有重要的功能，在平台生态系统中的身份验证、资源匹配、秩序治理的过程中发挥基础作用。〔33〕平台在信息处理过程中存在两个方面的难题。一是技术能力上的不足，存在处理准确程度不够的问题，亦即可能会出现误判、误伤等情况，这一问题在技术能力相对较弱的中小平台尤为明显。二是制度上的障碍，平台处理用户画像的行为，同样面临

〔31〕 参见邓丽：《最有利于未成年人原则的实践基础与制度理性》，载《政治与法律》2024 年第 6 期。

〔32〕 参见刘晓春：《平台信任机制构建中的调控型治理转向——以直播带货与〈广告法〉适用冲突为视角》，载《行政法学研究》2025 年第 2 期。

〔33〕 参见胡凌：《功能视角下个人信息的公共性及其实现》，载《法制与社会发展》2021 年第 5 期。

着现有制度上的两难境地。在未识别用户的情况下，平台主动对未成年人行为数据进行分析并用于识别，无法满足个人信息保护制度的合规要求，即未成年人个人信息的处理需要经过监护人的同意，并采取特殊保护措施等。〔34〕但是如果不进行此类识别，又会面临无法履行保护未成年人的法定义务的困难，进而面临更大的信息不足的困境。〔35〕在已识别未成年人用户的情况下，一方面对于未成年人用户画像的处理存在应用范围的限制，例如算法推荐方面的规范和限制，另一方面，在限制型未成年人模式下，大部分平台也并无激励而就模式内有限的内容及未成年人有限的使用行为开发专门的用户画像和匹配算法，从而导致在未成年人模式下缺乏有效的信息供应。

信息输出指的是面向未成年人提供信息内容等服务。理想状态下，能够吸引未成年人并有利于其身心健康发展的信息输出，应当是符合其兴趣偏好并且不会对其构成侵害或不良影响的内容。但是，如前所述，在信息收集和信息处理上都面临有效信息不足的困境，这就导致信息输出无法根据未成年人特点匹配其兴趣，而只能通过一刀切式的限制内容供应来降低风险。可以说，限制性措施在现有未成年人模式下，是信息严重不足的困境下的无奈选择，无法通过更加高质量的信息输出来吸引未成年人并降低风险，就只能通过减少和限制信息输出来控制风险，其代价则是无法留住未成年人的使用行为，同时也进一步加剧了信息不足的困境。

从未成年人模式面临的信息困境的三个层次可以看出，造成困境的原因，既有各方主体缺乏提供信息的意愿和激励的问题，也有法律制度的障碍，突出的如个人信息保护与未成年人识别之间的合规冲突。在整个模式运行缺乏有效信息支持的情况下，对于未成年人网络使用和保护的实际效果进行外部评估和评价，也同样面临有效信息不足的难题，政府和第三方机构很难对平台履行未成年人保护义务的成效给出具有信服力的评估结论。评估难题带来的是平台缺乏进一步采取有效措施的激励，从而进一步陷入限制型模式下信息困境的循环。

四、人工智能应用中未成年人保护模式的调控型转向

限制型未成年人保护模式面临的信息困境，导致实际的保护效果不尽理想，很难在控制和降低风险的同时为未成年人提供符合其兴趣偏好并有利于其健康发展的具有吸引力的内容和服务。在人工智能技术与应用蓬勃发展的阶段，限制型未成年人保护模式的局限性更为明显地凸现出来。如果继续沿用各种限制规则，不仅无法解决原有的问题，也与人工智能应用的特征和模式存在底层冲突。因此，有必要对限制型未成年人保护模式进行理念转型和机制改造，通过增加有效信息供给来破除信息不足的困境，并基于信息的充分获取、处理、输出，建立相应的调控型保护模式，为政府、平台、家长配置相应的保护和治理义务，确立相应的制度保障。

（一）人工智能应用中信息处理方式和能力的变化

在人工智能应用业态开启的人工智能时代，随着人工智能对信息处理、总结、生成能力的大幅提升，以及高度个性化、交互性的信息处理方式的发展，未成年人保护模式的信息不足困境，

〔34〕 参见《中华人民共和国个人信息保护法》第28、31条；《未成年人网络保护条例》第四章。

〔35〕 参见刘晓春：《人工智能时代用户画像的功能与治理：以未成年人保护为视角》，载《网络法律评论》第26卷，知识产权出版社2025年版，第49－50页。

在技术条件和商业模式上获得了突破的可能性。

在信息输入、信息处理、信息输出各层面，人工智能技术在理解、生成、推理、决策等方面能力的快速提升，都提供了相对于以往更高的信息处理能力、更低的信息处理成本以及更加精准的个性化信息处理方式。人工智能技术的信息处理方式和能力的变化，为未成年人保护模式从限制型转向调控型提供了技术上的可行性保障。

提供人工智能应用服务的平台，很大程度上将服务从公开领域转入私密领域，尽管这在一定程度上限制了其从公开内容进行风险监测和研判的能力，但是却基于其高度个性化、交互性的服务，取得了对于特定用户进行更加精准了解和分析的能力。平台基于人工智能应用的个性化调控能力，首先体现在比较精准的识别未成年人身份上，在大量个性化互动的前提下，平台至少能够比其他应用更有能力作出较为准确的用户画像，从而在不需要额外承担过高成本的前提下识别出未成年人身份。其次，平台在个性化的服务中，可以在使未成年人不产生“隔离”“限制”等体验的前提下，采取相应的风险防护措施，例如减少或者消除对于特定搜索结果的展示、特定内容和智能体的生成，并对可能存在的风险进行个性化提示。最后，平台亦有更强能力向未成年人用户提供个性化内容和服务，以拓展未成年人可以接触的高质量内容范围，来替代未成年人模式下通过限制措施来确保安全的思路。

对平台而言，对于提供此类个性化调控式的保护，也存在较高的激励。一方面，这种保护模式能够为未成年人提供较好的使用体验，从而有望增强他们的使用黏性，为平台培育高质量用户群体，这与平台的商业利益高度契合。另一方面，调控型保护实际上可以整合在人工智能应用的整体个性化服务之中，符合其运行逻辑并能够通过数据反馈进一步优化其服务，因此相比于孤立的未成年人模式更容易整合进平台的整体运营而无需过高的单独合规成本。信息收集、处理、输出的调控式保护过程，能够打破限制型保护模式中的信息困境，形成信息获取、整合、反馈的良性循环，基于充分的信息来实现保护未成年人和促进其发展并重的制度目标。

对于家长而言，他们事实上掌握了针对单个未成年人的重要信息优势，能够有助于突破信息困境，但由于种种原因并没有充分发挥其信息优势。尽管人工智能应用作为新技术新业态，依然持续对家长的监护能力提升速度构成挑战，但是对于拥有较高保护激励和较低保护成本的家长而言，如果能够提供操作简便、功能高效且不会对亲子关系构成威胁的管理工具，实际上很可能起到事半功倍的效果。在这个意义上，依托人工智能的个性化、定制化、生成式能力，将未成年人模式改造为调控型的面向家长的个性化赋能工具，促使其加入提供高质有效信息、破除信息困境的过程中，亦存在十分开阔的适用空间。

对于政府而言，尽管其在掌握个性化信息方面并不存在优势，但是在整体风险类型的研判和总结以及建立标准化的风险评估和治理举措方面，政府具有提供治理相关信息和指引的较强能力。政府从行业整体发展出发提供的风险和治理信息，正好可以弥补人工智能应用强调个性化信息处理的不足，通过将类型化、场景化、案例化的风险信息整合到人工智能信息处理过程中，为个性化的风险防范和未成年人保护提供更加精准、即时、高效的决策基础。

（二）调控型保护的构成层次

人工智能应用中调控型保护模式，与强调隔离、限制、人群类型化的未成年人模式相比，强

调精准化识别、个性化设置、赋能化发展等特征，基于人工智能高度个性化互动的服务特点和能力，探索面向人工智能时代的未成年人保护新模式。

在身份识别层面，平台基于个性化、互动性的人工智能应用服务，在取得用户有效授权基础上，应当进行特定维度用户画像的分析，建立未成年人身份识别技术模型和评估指标，结合大数据反映的行为特征和规律，对于用户是否符合未成年人用户特征进行研判和评估，形成初步评分。对于未成年人用户使用特征符合程度较高的账号应当启动相应的特殊设置和保护措施。平台也可以针对选定的用户进行提示，允许其通过提交更强的身份认证信息的方式排除账号特殊设置的可能性，例如要求手持身份证拍照或者通过人脸验证等。当然，在身份识别层面，平台需要根据个人信息保护等法律要求建立相应的合规流程，对于事实上存在的合规难点则需要在制度上给出合法性空间。〔36〕

在风险防范层面，平台可以基于识别出来的未成年人用户身份及其具体用户画像，采取个性化的风险防范策略和对应措施。原有未成年人模式下，或者将全部未成年人整体视为一个类型，或者进行大致的分龄化区分，都过于简单，而且年龄本身也只能是一个十分粗糙的分类标准，很难反映具体认知能力和风险防范能力的区别。人工智能应用场景下，用户画像的颗粒度可以大幅细化，并为不同未成年人用户的典型风险场景区分提供了可能性。例如，偏好游戏、直播、社交等服务类型的未成年人用户，可能面临不同类型和程度的风险场景。平台也可以建立统一的高风险场景监测和预警机制，与政府建立起风险信息共享系统，嵌入跨平台风险动态预警模块，对于网络性侵、隔空猥亵等高发的高风险场景进行有效提醒、警示，并在必要时对于用户权限实施限制，实现个性化精准化的风险预防和拦截效果。

在发展促进层面，平台应当发挥人工智能个性化服务和匹配优质内容的优势，在降低有害和不良内容传播和影响的同时，更要注重通过提供高质量内容来促进未成年人的教育和发展。特别是可以结合教学规律，根据未成年人主动发起的需求及其兴趣，“授之以渔”，通过引导而非禁止、限制，来提升未成年人识别、筛选、判断内容和行为优劣的数字素养。这些要求对于平台来说固然存在难度，特别是实力和规模有限的平台有可能无法承担专门开发此类技术模块的能力，但是，人工智能应用在教育方面的潜力可以为此提供乐观前景，并且可以通过培育专门开发此类功能的第三方服务市场来逐步实现。

在家长赋能层面，调控型保护模式下，一方面家长对未成年人账户的了解和管理可以通过间接、缓和的方式来实现，另一方面，家长对于未成年人保护的操作方式也可以更加简捷便利。对于家长的信息了解需求，人工智能应用不必披露原始互动数据，而可以通过生成抽象层次较高的用户使用报告来实现，而且报告可以根据多个维度进行定制，也可以根据家长的具体需求、未成年人的具体行为特征进行个性化展示。家长也极有可能通过简单的对话式指令，对于未成年人账号的保护措施进行设置，而无须学习和探索繁复的预设功能。例如，对于困扰家长的未成年人大额消费行为，家长可以指令形式直接要求平台打开刷脸支付等验证要求，并对未成年人改动设置

〔36〕 参见刘晓春：《人工智能时代用户画像的功能与治理：以未成年人保护为视角》，载《网络法律评论》第 26 卷，知识产权出版社 2025 年版，第 58－59 页。

的权限进行限制。在前沿技术的支持下，人工智能应用平台应有能力通过接受自然语言指令来修改各种个性化设置。实现这一保护效果的技术能力，实际上也可以从人工智能应用，拓展到几乎所有需要展开未成年人保护的网络服务领域。

（三）调控型未成年人保护模式的制度保障

从限制型未成年人保护模式转向调控型模式，核心问题在于破除限制型模式下的信息困境，改变通过限制信息输入输出来控制风险的方式，充分激发人工智能技术在信息处理方面的优势，形成信息流动的良性循环。理念和机制的转变，需要在法律制度上作出积极调整，一方面为信息的获取、处理和流动解除制度障碍，另一方面也要在各方主体之间根据其信息优势进行相应的权责分配并提供激励，使得调控型未成年人保护模式能够在各方主体有效协同之下，更加充分实现保护与发展并举的目标。

首先，从法律规则上，需要确立基于保护为特定目的的未成年人个人信息和用户画像处理行为的合法性基础。未成年人个人信息在我国个人信息保护制度中作为敏感个人信息进行界定和保护，对其进行处理需要遵循严格的同意规则和其他保护措施。在限制型未成年人保护模式中，信息困境的一个重要制度成因，就是严格的未成年人个人信息合规义务与未成年人保护义务之间的冲突和悖论。破除信息困境，从限制型转向调控型模式，重要的制度前提就是，需要针对未成年人个人信息和用户画像处理行为，确立以识别和保护为目的的同意例外规则，并针对未成年人用户画像行为确立较为明确的正向指引规则。由此改变涉及未成年人个人信息处理就“谈虎色变”的情况，为平台积极展开调控型保护模式的建设提供制度上的保障和激励。当然，这些基于识别和保护目的的未成年人个人信息和用户画像处理行为也应当符合目的限定的规范，建立相应的评估和检测机制，防范超越目的的信息处理行为，更要严格治理利用未成年人用户画像实施有损其身心健康的滥用行为。

其次，政府和平台之间建立新型技术条件下的“发包”和沟通机制，在调控型模式中有效发挥其各自的信息优势，共同建构协同治理机制。〔37〕

在调控型未成年人保护模式之下，政府基于其在治理信息方面的优势和能力，可以主要通过三种方式来展开治理。一是确立宏观原则和机制，在立法层面建立体系化框架，除了基础立法，也体现在具体领域治理的相关规定中，如内容、算法、网络暴力、饭圈治理等等。〔38〕二是定期发送微观指令，针对新现象新问题新风险开展专项行动，整治重点问题，典型实践如网信部门长期开展的“清朗”行动未成年人保护、暑期春节专项等。〔39〕三是进行中间层次的规则建设，即介于前述宏观和微观模式之间，基于微观实践定期总结典型案例和对策，制定重点领域的规范性文件、标准、指引，采用“软硬兼施”的措施，指导平台建立动态风险治理模块及其细则。〔40〕

---

〔37〕 参见胡凌：《平台发包制：当代中国平台治理的内在逻辑》，载《文化纵横》2023 年第 4 期。

〔38〕《网络信息内容生态治理规定》《互联网信息服务算法推荐管理规定》《网络暴力信息治理规定》等网络空间治理领域的相关规定，其中都针对未成年人保护设有专门条款，并要求对于未成年人实施特殊和优先保护。

〔39〕 参见中国网络社会组织联合会：《关于未成年人，这一重磅报告发布!》，载 https://mp.weixin.qq.com/s/13n8cn3-wMDEmq1bIcrPlw，最后访问时间：2024 年 12 月 20 日。

〔40〕 网信部门通过总结典型案例进行指导的实践，参见《网信部门从严打击网上侵害未成年人合法权益行为》，载 https://www.cac.gov.cn/2024-12/19/c_1736308605590121.htm，最后访问时间：2024 年 12 月 20 日。

平台作为具有最有能力获取信息优势的治理主体，在调控式未成年人保护模式方面的“主体责任”，可以从三个层面来建构。首先，平台需要建立完备的治理机制体系，在信息内容建设和管理、个人信息保护、网络沉迷防范、算法推送、网络暴力以及其他侵害预防等领域，结合不同业态特点建立未成年人保护的特殊和优先机制，而具有枢纽性和整合性的人工智能应用服务更需要建立综合性的防范机制。〔41〕并且，对未成年人具有实质重大影响力的平台，需要建立更加严格的保护和防范体系，并定期展开内外部的专业评估。〔42〕其次，平台需要建设未成年人保护的基础设施，并开发相应的技术模块和功能，在信息输入、处理、输出阶段分别建设技术能力，实现保护的现实可操作性。最后，平台对未成年人网络保护的“主体责任”体现为一种兜底责任，在高度动态化的数字环境下，即使缺乏来自政府或者外界的明确指令，平台也需要自行展开创新举措的探索，积极采取保护措施，并在无法证明已经充分履责的情况下承担保护不力带来的声誉受损后果。〔43〕

再次，需要建立起促进和激励家长参与调控型治理的制度框架。家长作为能够以较低成本提供有效信息的主体，有必要在调控型保护模式中承担更多义务。除了通过提高未成年人保护模式的使用便捷程度来优化家长使用体验之外，也需要在制度上建立具有积极导向的规则，使家长在未成年人行为和后果方面承担相应的义务和责任。例如，在未成年人未经家长同意进行大额充值打赏的领域，应当由家长承担消费行为相关的注意义务，家长未能尽到注意义务的，需要为其过错承担相应的法律责任，从而为其提升监护和管理意识提供必要的激励和保障。

最后，对于调控型未成年人模式，应当建立起循序渐进的评估机制，并给出容错空间。无论是平台在信息获取和处理中合规义务的履行，技术模型处理的准确性程度，还是调控型保护中的个性化信息输出和保护效果，都应当对其保持适度的宽容，允许存在平台探索和创新的空间。从限制型的信息困境转向基于充分个性化处理的调控型模式，存在着技术、管理、规范、评估等各个环节的难点和不确定性。只有秉持包容和开放的容错态度，为平台的自治性探索留出充分的空间，才能为调控型模式的转型和发展提供充分的激励，并有望在此基础上拓展应用空间，将人工智能技术带来的治理前景传递到更加广阔的网络空间。

五、结　论

人工智能时代背景下，未成年人使用人工智能新型应用的行为具有高度交互性、私密性、工具性和枢纽性等特点，从而对未成年人这一特殊群体的保护提出了新的问题和挑战。由于家庭监护不能单独实现对网络和人工智能时代未成年人的充分保护，政府和互联网平台都承担起了相应的保护职责，形成“政府—平台—家长”三方为主的保护和治理机制。未成年人模式作为具有中

〔41〕 参见林维、刘晓春：《中国未成年人网络保护发展现状与展望》，载林维主编：《未成年人网络保护发展报告（2021）》，中国社会科学出版社 2022 年版，第 8－10 页。

〔42〕 参见《未成年人网络保护条例》第 20 条；唐赟：《未成年人网络保护的平台责任及其实现——兼论〈未成年人网络保护条例〉第 20 条的落实》，载《少年儿童研究》2024 年第 5 期。

〔43〕 参见戴昕：《平台责任与社会信任》，载《法律科学（西北政法大学学报）》2023 年第 2 期。

国特色的保护机制，已经取得较为成熟的实践成果，但在实施效果上依然面临质疑。

我国未成年人模式的运行机理包括身份识别、权限设置和家长赋能三个方面。然而，由于具有明显的限制型特征，该模式面临未成年人保护层面信息输入、处理、输出三个层次的信息困境，导致保护要求难以落地。人工智能时代，限制型未成年人模式无法满足未成年人能力发展的现实需求。基于人工智能技术带来的平台识别和保护能力的提升，需要将未成年人模式进行改造，转向调控型保护模式，即基于更加充分的信息输入、处理和输出，破除信息困境，突出个性化保护和发展，从身份识别、风险防范、发展促进、家长赋能四个层次展开系统性重塑，充分调动家长使用保护模式的积极性，并建构保障性的制度规则，以个性化、精准化、赋能化的方式，促进未成年人保护和发展目标的均衡实现。

Abstract: Minors' use of artificial intelligence applications is characterized by high interactivity, privacy, instrumentality, and hub-like nature, posing new challenges to the protection of minors. China's minor protection mode has achieved practical results, but its implementation effect is still questioned. The operational mechanism of China's minor protection mode includes three aspects: identity identification, permission setting, and parental empowerment, which has obvious "restriction-oriented" characteristics. It faces information dilemmas at three levels: input, processing, and output, making it difficult to implement effective protection. In the context of artificial intelligence, it is necessary to transform the minor protection mode into a modulatory protection mode. Based on the improved information processing capability and reduced costs brought by artificial intelligence technology, we should break through information dilemmas, carry out systematic reconstruction from four levels: identity identification, risk prevention, development promotion, and parental empowerment, and establish institutional guarantees to realize modulatory protection for minors in a personalized, precise, and empowering manner.

Key Words: minor online protection, minor mode, artificial intelligence, modulatory protection

(责任编辑：张金平)