（2025 年 10 月）

目录

一、法规速递

《政务领域人工智能大模型部署应用指引》

发文机关：国家互联网信息办公室

发布时间：2025.10.10

生效时间：2025.10.10

为深入贯彻落实党中央、国务院决策部署，规范和引导人工智能大模型在政务领域的发展与应用，提升政务数字化智能化治理和服务水平，制定本指引。本指引主要为各级政务部门提供人工智能大模型部署应用的工作导向和基本参照，将根据实践进展，结合人工智能大模型发展和应用的新形势、新要求，进行动态调整。

一、总体要求

坚持以习近平新时代中国特色社会主义思想为指导，深入贯彻落实党的二十大和二十届二中、三中全会精神，全面贯彻习近平总书记关于网络强国的重要思想，完整准确全面贯彻新发展理念，统筹高质量发展和高水平安全，坚持系统谋划、集约发展，以人为本、规范应用，共建共享、高效协同，安全稳妥、务求实效，有序推进人工智能大模型技术、产品和服务在政务领域的部署、应用和持续优化，充分发挥人工智能大模型在复杂语义理解与推理、多模态内容生成、知识整合与分析等方面的优势，为工作人员提供高效辅助，为公众和企业提供便捷服务，推动政务创新发展，提升治理效能、优化服务管理、辅助科学决策。

二、应用场景

政务部门可围绕政务服务、社会治理、机关办公和辅助决策等工作中的共性、高频需求，因地制宜、结合实际，选择典型场景进行人工智能大模型探索应用。主要包括以下参考场景：

（一）政务服务类

1.智能问答。整合本地区、本部门、本领域业务资源和知识库等数据，利用自然语言理解、检索增强生成和知识图谱等技术，提供便捷的在线政务咨询服务，加强对公众诉求的准确理解，实时生成参考回答，帮助解决公众疑惑，提升信息获取便捷性。

2.辅助办理。整合政务服务办事指南、常见问题、用户评价和历史办理记录等数据，利用智能匹配和自动化处理等技术，提供智能导办、个性引导、表单预填、辅助审核、进度查询和提醒等一站式政务辅助办理服务，辅助工作人员高效审核材料，支撑公众和企业便捷办理事项。

3.政策服务直达快享。构建政策服务知识库，细化政策要求、政策标签、推送条件、申兑流程等相关内容，利用“政策找人”“政策找企业”算法模型，加强公众和企业需求分析，实现政策智能匹配，推进惠民利民、惠企利企服务主动精准送达和一站式办理。

（二）社会治理类

4.智能监测巡检。利用无人机、视频监控、智能传感器等设备和计算机视觉等技术，对监控视频、图像、物联感知数据等进行实时分析，辅助工作人员实时监测房屋、道路、燃气、桥梁、供水、排水、供热、综合管廊等基础设施，及时发现异常行为、环境问题或设施故障等，自动识别潜在风险隐患，及时进行提醒，并根据异常情况和严重程度提供处置建议，提高监测巡检效率。

5.辅助执法监管。采用语音识别、视频分析、知识图谱、逻辑推理等技术，辅助执法人员将案件信息实时录入系统、穿透式发现问题线索、生成案件报告、快速检索法律依据和司法解释、查询类似典型案例等，提供针对性案件办理建议，提高执法监管效率和规范性。

6.市场风险预测。运用生成式时间序列分析模型和异常检测算法等，对各类市场数据进行监测和深度分析，捕捉市场动向，包括经济指标波动、异常情况等，预测可能出现的市场风险，研判对经济社会带来的影响和经济走势，并及时发出预警，为政府管理和社会治理提供支撑。

（三）机关办公类

7.辅助文书起草。利用大语言模型的生成能力，通过构建本地知识库和预设模板，为工作人员提供写作建议、辅助起草文书，对格式和内容等进行检查、校对和优化，提高工作效率，减轻基层负担。

8.资料检索。利用知识图谱构建和信息检索等技术，准确理解工作人员资料检索需求，实现政务信息快速检索、精准定位、多维度排序、智能关联和对比分析等，帮助工作人员提升资料检索效率和准确性。

9.智能分办。利用自然语言理解和多模态识别等技术，构建多维度任务分类和分办规则，对来文、来电、工单等任务进行自动识别、准确分类、辅助填写和优先级排序，实现辅助分发和智能派单，提高任务分办效率。

（四）辅助决策类

10.灾害预警。对来自卫星、地面传感器、地质监测站，以及预报预警、灾害风险普查等多源、多维、多模态数据进行大数据关联和综合分析研判，识别异常波动情况，预测可能发生的自然灾害，并提前发出预警，辅助政务部门及时采取有效措施，减轻灾害风险，减少灾害损失。

11.应急处置。利用强化学习等技术，对社会公共安全等突发事件的性质、特点、危害程度、影响范围、发展趋势和公众反应等进行分析研判，及时发现和预警风险隐患，基于突发场景、力量资源分布等快速模拟应急处置方案效果，提供科学合理的应急处置建议，优化救援资源配置，提高应急响应速度和效率。

12.政策评估。利用人工智能大模型推断分析能力与数据挖掘能力，对公众反馈、市场反应、经济指标和社会满意度等进行分析，构建多维度指标，评估政策目标实现程度、政策影响力和潜在问题，支撑政策制定部门进行政策优化。

13.智能辅助评审。利用自学习泛化认识、类人化评审推理、多模态智能解析等能力，对照有关要求开展项目评审，对项目文件内容进行深度扫描、智能解析，提出评审意见建议，辅助提高项目评审效率和科学性。

三、规范部署

政务部门应结合工作实际和场景特点，充分论证人工智能大模型的应用需求、实施路径、功能设计等，选择适宜的部署模式，统筹推进实施，推动共建共享，提升建设管理效能。

（一）合理选择实施路径

政务部门应根据不同政务场景需求与现有技术基础，审慎选择人工智能大模型实施路径。对于智能问答、辅助文书起草等通用性较强、数据资源丰富的场景，需采用市场上成熟，

并已完成网信部门备案的模型产品和服务。对于辅助执法监管、市场风险预测等专业性较强、业务逻辑复杂的场景，可利用领域专家知识和专业数据进行针对性训练，打造垂直模型。在保障安全和不泄露国家秘密、工作秘密和敏感信息等的前提下，充分利用互联网算力和模型资源，开展政务领域人工智能大模型部署应用。鼓励探索政务智能体、具身智能等创新应用。

（二）统筹集约开展部署

政务部门应以统筹集约的方式开展政务领域人工智能大模型部署，依托“东数西算”和全国一体化算力网，统筹推进智能算力基础设施布局，并实施集中统一的安全管理和体系化技术防护措施，避免“碎片化”安全风险。有条件的中央和国家机关部门、省（自治区、直辖市）可统一部署智能算力资源、人工智能大模型，面向下属单位或下辖地区提供电子政务外网环境下的人工智能大模型服务。地市应在省（自治区、直辖市）统一要求下开展部署应用，县级及以下原则上应复用上级的智能算力和模型资源开展应用和服务，不再独立进行政务大模型建设和部署。

（三）探索实现统管复用

政务部门应探索构建“一地建设、多地多部门复用”的集约化部署模式，统筹推进政务大模型部署应用，防止形成“模型孤岛”。省（自治区、直辖市）应搭建政务领域人工智能大模型统一服务平台，并与政务云管理平台、政务应用和组件管理平台等融合共建，将区域内电子政务外网智能算力、政务大模型、政务数据集等资源纳入统一管理，形成要素资源“一本账”，支撑政务大模型运行监测，提供资源申请与调度服务，推动高效复用。国家行业主管部门按照业务需求和发展需要探索细分领域政务垂直大模型的统一训练与构建，加强与省（自治区、直辖市）协同部署，深化跨层级、跨地域的行业领域智能化赋能。垂直管理部门应强化模型、算力、数据等资源的统筹部署和管理，避免资源浪费。

（四）持续夯实数据基础

政务部门应加强政务数据治理，持续提升数据质量，加快构建客观反映公共政策、制度规范、业务流程和治理实效的高质量政务数据集和知识库，支撑政务大模型的优化训练。分类分级管理政务大模型涉及数据，加强训练数据、微调数据、知识库等管理，建立台账并详细记录数据来源、类型和规模等信息，确保数据来源可靠可追溯、内容准确有效。依托政务数据共享协调机制，统筹数据治理成果，推进高质量政务数据集的共建共享和生成数据的归集治理。探索基于大模型的政务知识治理路径，打造可信知识库，确保数据源的权威性、准确性和时效性。

四、运行管理

政务部门应强化政务领域人工智能大模型运行管理，健全管理制度、运行模式和安全要求，有序推进人工智能大模型技术、产品和服务在政务领域部署应用。

（一）明确应用管理要求

政务部门应统筹减负和赋能，严格落实《整治形式主义为基层减负若干规定》《关于防治“指尖上的形式主义”的若干意见》等相关要求，避免盲目追求技术领先、概念创新，避免重复建设、无效建设，避免未审先建、建而不管，避免强制使用、无效使用，避免数据多头采集、重复索要，切实防范“数字形式主义”。中央和国家机关政务领域人工智能大模型部署应用应纳入国家政务信息化规划统筹。政务部门应建立健全涵盖政务领域人工智能大模型部署应用全周期管理体系，明确应用方式和边界，落实人工智能大模型“辅助型”定位，及时解决部署与应用过程中出现的新问题。应在政务大模型应用界面显著位置设置风险提示，明确告知大模型服务的局限性，做好大模型输出内容标识。对于智能问答、辅助办理等代表政务部门面向公众和企业提供服务的人工智能大模型应用场

景，应严格执行内容审核制度流程，结合场景特点和技术能力合理采用人工审核、生成内容实时风控、多模型交叉校验等措施，防范模型“幻觉”等风险，确保输出内容不超出业务范围，保障内容准确性，维护政务部门公信力。

（二）持续推动迭代优化

政务部门应将持续迭代优化作为人工智能大模型部署应用的关键环节，建立常态化更新机制，加快功能优化，深化场景应用。密切关注技术发展动态，持续更新优化政务领域人工智能大模型的基础模型和安全能力。建立高效数据收集处理机制，及时更新支撑人工智能大模型运行的输入数据和知识库，并适时进行清洗、标注，补充优化训练数据集，持续提升模型能力。建立政务领域人工智能大模型用户评价反馈机制，及时收集、处理用户需求，以用户反馈驱动迭代优化。

（三）扎实做好安全管理

政务部门应建立安全责任制度，明确数据处理、大模型训练和场景应用各阶段参与主体的安全职责和任务，做好用户身份识别和权限管理。政务部门提供人工智能大模型服务时，应遵守《生成式人工智能服务管理暂行办法》等相关规定，使用具有合法来源的数据和基础模型，依法履行算法备案和安全评估等义务，与使用者签订服务协议，明确双方权利义务。构建政务领域人工智能大模型分类分级治理制度，完善安全管理流程，针对可能出现的安全风险，制定应急处置预案。做好政务大模型对抗攻击的检测与处置，识别并拦截提示词注入、资源消耗攻击等。加强政务大模型内容安全管理，综合运用语义识别、规则库、模型算法等，做好多模态输入输出内容的识别、分析与管控，建立合理的代答、拒答机制，及时发现并处置违法和不良信息、敏感内容等。发挥新闻媒体内容审核优势，做好政务大模型训练数据的内容审核把关，加强政务大模型内容监测管理。做好政务大模型应用运行日志管理，定期对日志记录进行审计。推动形成安全风险威胁信息共享和应急处置机制，按照规定及时处置并报告安全事件，提升人工智能安全风险

应对能力。

（四）严格落实保密要求

政务部门在模型训练、部署应用等过程中应加强数据安全保密和个人信息保护，坚持底线思维，严格落实“涉密不上网、上网不涉密”等保密纪律要求，采取加装保密“护栏”等措施，防止国家秘密、工作秘密和敏感信息等输入非涉密人工智能大模型，防范敏感数据汇聚、关联引发的泄密风险。制定完善人工智能大模型在政务领域应用相关保密管理制度，规范人工智能大模型选型、部署、训练、使用、废止等全流程保密管理。涉密信息系统应用人工智能大模型按照国家保密行政管理部门要求稳妥推进。

五、保障措施

（一）加强组织实施

加强统筹协调，稳妥有序推动人工智能大模型在政务服务、社会治理、机关办公、辅助决策等领域规范应用。加快推进政务领域人工智能大模型国家标准体系建设和重点标准研制，明确应用效果评估、系统技术要求、智能技术应用等工作规范，支撑部署应用取得实效。及时总结推广政务领域部署应用人工智能大模型的典型场景和创新应用，推动复用增效。加强政务领域人工智能大模型部署应用经费保障，引入市场化的产品和服务竞争机制，探索企业建设运营、政府购买服务、按使用情况结算费用的运作模式，营造高效、可持续的政务大模型生态。

（二）开展监测评估

构建政务领域人工智能大模型部署应用全流程监测评估体系，适时开展监测评估工作。建立政务大模型安全测评机制，上线前对模型算法、生成内容、应用功能、配置环境、挂接数据、漏洞风险等进行充分测试验证，对发现的问题隐患进行整改加固。加强政务领域人工智能大模型系统运行状态、响应时间、准确性、安全性和潜在风险的实时监测

分析，及时发现问题，并采取有效措施解决。做好人工智能大模型应用效能评价，及时总结经验，持续迭代优化，推动部署应用取得实效。

（三）做好培训宣传

开发涵盖人工智能大模型理论、技术、应用、安全、伦理、产业等内容的培训课程体系，开展人工智能素养和技能培训，提升领导干部对人工智能的认知水平，增强工作人员应用能力和水平。面向公众做好宣传教育，提升全民数字素养，积极回应用户关切，正确引导社会对政务领域人工智能大模型适用人群、场景、用途的认识和预期，客观反映人工智能大模型在优化政务服务、满足公众和企业需求、提升社会治理水平等方面的作用。

《中华人民共和国网络安全法（2025 修正）》

发文机关：全国人民代表大会常务委员会

发布时间：2025.10.28

生效时间：2026.01.01

（2016 年 11 月 7 日第十二届全国人民代表大会常务委员会第二十四次会议通过　根据 2025 年 10 月 28 日第十四届全国人民代表大会常务委员会第十八次会议《关于修改〈中华人民共和国网络安全法〉的决定》修正）

第一章 总 则

第一条　为了保障网络安全，维护网络空间主权和国家安全、社会公共利益，保护公民、法人和其他组织的合法权益，促进经济社会信息化健康发展，制定本法。

第二条　在中华人民共和国境内建设、运营、维护和使用网络，以及网络安全的监督管理，适用本法。

第三条　网络安全工作坚持中国共产党的领导，贯彻总体国家安全观，统筹发展和安全，推进网络强国建设。

第四条　国家坚持网络安全与信息化发展并重，遵循积极利用、科学发展、依法管理、确保安全的方针，推进网络基础设施建设和互联互通，鼓励网络技术创新和应用，支持培养网络安全人才，建立健全网络安全保障体系，提高网络安全保护能力。

第五条　国家制定并不断完善网络安全战略，明确保障网络安全的基本要求和主要目标，提出重点领域的网络安全政策、工作任务和措施。

第六条　国家采取措施，监测、防御、处置来源于中华人民共和国境内外的网络安全风险和威胁，保护关键信息基础设施免受攻击、侵入、干扰和破坏，依法惩治网络违法犯罪活动，维护网络空间安全和秩序。

第七条　国家倡导诚实守信、健康文明的网络行为，推动传播社会主义核心价值观，采取措施提高全社会的网络安全意识和水平，形成全社会共同参与促进网络安全的良好环境。

第八条　国家积极开展网络空间治理、网络技术研发和标准制定、打击网络违法犯罪等方面的国际交流与合作，推动构建和平、安全、开放、合作的网络空间，建立多边、民主、透明的网络治理体系。

第九条　国家网信部门负责统筹协调网络安全工作和相关监督管理工作。国务院电信主管部门、公安部门和其他有关机关依照本法和有关法律、行政法规的规定，在各自职责范围内负责网络安全保护和监督管理工作。

县级以上地方人民政府有关部门的网络安全保护和监督管理职责，按照国家有关规定确定。

第十条　网络运营者开展经营和服务活动，必须遵守法律、行政法规，尊重社会公德，遵守商业道德，诚实信用，履行网络安全保护义务，接受政府和社会的监督，承担社会责任。

第十一条　建设、运营网络或者通过网络提供服务，应当依照法律、行政法规的规定和国家标准的强制性要求，采取技术措施和其他必要措施，保障网络安全、稳定运行，有效应对网络安全事件，防范网络违法犯罪活动，维护网络数据的完整性、保密性和可

用性。

第十二条　网络相关行业组织按照章程，加强行业自律，制定网络安全行为规范，指导会员加强网络安全保护，提高网络安全保护水平，促进行业健康发展。

第十三条　国家保护公民、法人和其他组织依法使用网络的权利，促进网络接入普及，提升网络服务水平，为社会提供安全、便利的网络服务，保障网络信息依法有序自由流动。

任何个人和组织使用网络应当遵守宪法法律，遵守公共秩序，尊重社会公德，不得危害网络安全，不得利用网络从事危害国家安全、荣誉和利益，煽动颠覆国家政权、推翻社会主义制度，煽动分裂国家、破坏国家统一，宣扬恐怖主义、极端主义，宣扬民族仇恨、民族歧视，传播暴力、淫秽色情信息，编造、传播虚假信息扰乱经济秩序和社会秩序，以及侵害他人名誉、隐私、知识产权和其他合法权益等活动。

第十四条　国家支持研究开发有利于未成年人健康成长的网络产品和服务，依法惩治利用网络从事危害未成年人身心健康的活动，为未成年人提供安全、健康的网络环境。

第十五条　任何个人和组织有权对危害网络安全的行为向网信、电信、公安等部门举报。收到举报的部门应当及时依法作出处理；不属于本部门职责的，应当及时移送有权处理的部门。

有关部门应当对举报人的相关信息予以保密，保护举报人的合法权益。

第二章 网络安全支持与促进

第十六条　国家建立和完善网络安全标准体系。国务院标准化行政主管部门和国务

院其他有关部门根据各自的职责，组织制定并适时修订有关网络安全管理以及网络产品、服务和运行安全的国家标准、行业标准。

国家支持企业、研究机构、高等学校、网络相关行业组织参与网络安全国家标准、行业标准的制定。

第十七条　国务院和省、自治区、直辖市人民政府应当统筹规划，加大投入，扶持重点网络安全技术产业和项目，支持网络安全技术的研究开发和应用，推广安全可信的网络产品和服务，保护网络技术知识产权，支持企业、研究机构和高等学校等参与国家网络安全技术创新项目。

第十八条　国家推进网络安全社会化服务体系建设，鼓励有关企业、机构开展网络安全认证、检测和风险评估等安全服务。

第十九条　国家鼓励开发网络数据安全保护和利用技术，促进公共数据资源开放，推动技术创新和经济社会发展。

第二十条　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，促进人工智能应用和健康发展。

国家支持创新网络安全管理方式，运用人工智能等新技术，提升网络安全保护水平。

第二十一条　各级人民政府及其有关部门应当组织开展经常性的网络安全宣传教育，并指导、督促有关单位做好网络安全宣传教育工作。

大众传播媒介应当有针对性地面向社会进行网络安全宣传教育。

第二十二条　国家支持企业和高等学校、职业学校等教育培训机构开展网络安全相关教育与培训，采取多种方式培养网络安全人才，促进网络安全人才交流。

第三章 网络运行安全

第一节 一般规定

第二十三条　国家实行网络安全等级保护制度。网络运营者应当按照网络安全等级保护制度的要求，履行下列安全保护义务，保障网络免受干扰、破坏或者未经授权的访问，防止网络数据泄露或者被窃取、篡改：

（一）制定内部安全管理制度和操作规程，确定网络安全负责人，落实网络安全保护责任；

（二）采取防范计算机病毒和网络攻击、网络侵入等危害网络安全行为的技术措施；

（三）采取监测、记录网络运行状态、网络安全事件的技术措施，并按照规定留存相关的网络日志不少于六个月；

（四）采取数据分类、重要数据备份和加密等措施；

（五）法律、行政法规规定的其他义务。

第二十四条　网络产品、服务应当符合相关国家标准的强制性要求。网络产品、服务的提供者不得设置恶意程序；发现其网络产品、服务存在安全缺陷、漏洞等风险时，应当立即采取补救措施，按照规定及时告知用户并向有关主管部门报告。

网络产品、服务的提供者应当为其产品、服务持续提供安全维护；在规定或者当事人约定的期限内，不得终止提供安全维护。

网络产品、服务具有收集用户信息功能的，其提供者应当向用户明示并取得同意；涉及用户个人信息的，还应当遵守本法和有关法律、行政法规关于个人信息保护的规定。

第二十五条　网络关键设备和网络安全专用产品应当按照相关国家标准的强制性要求，由具备资格的机构安全认证合格或者安全检测符合要求后，方可销售或者提供。国家网信部门会同国务院有关部门制定、公布网络关键设备和网络安全专用产品目录，并推动安全认证和安全检测结果互认，避免重复认证、检测。

第二十六条　网络运营者为用户办理网络接入、域名注册服务，办理固定电话、移动电话等入网手续，或者为用户提供信息发布、即时通讯等服务，在与用户签订协议或者确认提供服务时，应当要求用户提供真实身份信息。用户不提供真实身份信息的，网络运营者不得为其提供相关服务。

国家实施网络可信身份战略，支持研究开发安全、方便的电子身份认证技术，推动不同电子身份认证之间的互认。

第二十七条　网络运营者应当制定网络安全事件应急预案，及时处置系统漏洞、计算机病毒、网络攻击、网络侵入等安全风险；在发生危害网络安全的事件时，立即启动应急预案，采取相应的补救措施，并按照规定向有关主管部门报告。

第二十八条　开展网络安全认证、检测、风险评估等活动，向社会发布系统漏洞、计算机病毒、网络攻击、网络侵入等网络安全信息，应当遵守国家有关规定。

第二十九条　任何个人和组织不得从事非法侵入他人网络、干扰他人网络正常功能、窃取网络数据等危害网络安全的活动；不得提供专门用于从事侵入网络、干扰网络正常功能及防护措施、窃取网络数据等危害网络安全活动的程序、工具；明知他人从事危害网络安全的活动的，不得为其提供技术支持、广告推广、支付结算等帮助。

第三十条　网络运营者应当为公安机关、国家安全机关依法维护国家安全和侦查犯罪的活动提供技术支持和协助。

第三十一条　国家支持网络运营者之间在网络安全信息收集、分析、通报和应急处置等方面进行合作，提高网络运营者的安全保障能力。

有关行业组织建立健全本行业的网络安全保护规范和协作机制，加强对网络安全风险的分析评估，定期向会员进行风险警示，支持、协助会员应对网络安全风险。

第三十二条　网信部门和有关部门在履行网络安全保护职责中获取的信息，只能用于维护网络安全的需要，不得用于其他用途。

第二节 关键信息基础设施的运行安全

第三十三条　国家对公共通信和信息服务、能源、交通、水利、金融、公共服务、电子政务等重要行业和领域，以及其他一旦遭到破坏、丧失功能或者数据泄露，可能严重危害国家安全、国计民生、公共利益的关键信息基础设施，在网络安全等级保护制度的基础上，实行重点保护。关键信息基础设施的具体范围和安全保护办法由国务院制定。

国家鼓励关键信息基础设施以外的网络运营者自愿参与关键信息基础设施保护体系。

第三十四条　按照国务院规定的职责分工，负责关键信息基础设施安全保护工作的部门分别编制并组织实施本行业、本领域的关键信息基础设施安全规划，指导和监督关键信息基础设施运行安全保护工作。

第三十五条　建设关键信息基础设施应当确保其具有支持业务稳定、持续运行的性能，并保证安全技术措施同步规划、同步建设、同步使用。

第三十六条　除本法第二十三条的规定外，关键信息基础设施的运营者还应当履行下列安全保护义务：

（一）设置专门安全管理机构和安全管理负责人，并对该负责人和关键岗位的人员进行安全背景审查；

（二）定期对从业人员进行网络安全教育、技术培训和技能考核；

（三）对重要系统和数据库进行容灾备份；

（四）制定网络安全事件应急预案，并定期进行演练；

（五）法律、行政法规规定的其他义务。

第三十七条　关键信息基础设施的运营者采购网络产品和服务，可能影响国家安全的，应当通过国家网信部门会同国务院有关部门组织的国家安全审查。

第三十八条　关键信息基础设施的运营者采购网络产品和服务，应当按照规定与提供者签订安全保密协议，明确安全和保密义务与责任。

第三十九条　关键信息基础设施的运营者在中华人民共和国境内运营中收集和产生的个人信息和重要数据应当在境内存储。因业务需要，确需向境外提供的，应当按照国家网信部门会同国务院有关部门制定的办法进行安全评估；法律、行政法规另有规定的，依照其规定。

第四十条　关键信息基础设施的运营者应当自行或者委托网络安全服务机构对其网络的安全性和可能存在的风险每年至少进行一次检测评估，并将检测评估情况和改进措施报送相关负责关键信息基础设施安全保护工作的部门。

第四十一条　国家网信部门应当统筹协调有关部门对关键信息基础设施的安全保护采取下列措施：

（一）对关键信息基础设施的安全风险进行抽查检测，提出改进措施，必要时可以委托网络安全服务机构对网络存在的安全风险进行检测评估；

（二）定期组织关键信息基础设施的运营者进行网络安全应急演练，提高应对网络安全事件的水平和协同配合能力；

（三）促进有关部门、关键信息基础设施的运营者以及有关研究机构、网络安全服务机构等之间的网络安全信息共享；

（四）对网络安全事件的应急处置与网络功能的恢复等，提供技术支持和协助。

第四章 网络信息安全

第四十二条　网络运营者应当对其收集的用户信息严格保密，并建立健全用户信息保护制度。

网络运营者处理个人信息，应当遵守本法和《中华人民共和国民法典》、《中华人民共和国个人信息保护法》等法律、行政法规的规定。

第四十三条　网络运营者收集、使用个人信息，应当遵循合法、正当、必要的原则，公开收集、使用规则，明示收集、使用信息的目的、方式和范围，并经被收集者同意。

网络运营者不得收集与其提供的服务无关的个人信息，不得违反法律、行政法规的规定和双方的约定收集、使用个人信息，并应当依照法律、行政法规的规定和与用户的约定，处理其保存的个人信息。

第四十四条　网络运营者不得泄露、篡改、毁损其收集的个人信息；未经被收集者同意，不得向他人提供个人信息。但是，经过处理无法识别特定个人且不能复原的除外。

网络运营者应当采取技术措施和其他必要措施，确保其收集的个人信息安全，防止信息泄露、毁损、丢失。在发生或者可能发生个人信息泄露、毁损、丢失的情况时，应当立即采取补救措施，按照规定及时告知用户并向有关主管部门报告。

第四十五条　个人发现网络运营者违反法律、行政法规的规定或者双方的约定收集、使用其个人信息的，有权要求网络运营者删除其个人信息；发现网络运营者收集、存储的其个人信息有错误的，有权要求网络运营者予以更正。网络运营者应当采取措施予以删除或者更正。

第四十六条　任何个人和组织不得窃取或者以其他非法方式获取个人信息，不得非法出售或者非法向他人提供个人信息。

第四十七条　依法负有网络安全监督管理职责的部门及其工作人员，必须对在履行

职责中知悉的个人信息、隐私和商业秘密严格保密，不得泄露、出售或者非法向他人提供。

第四十八条　任何个人和组织应当对其使用网络的行为负责，不得设立用于实施诈骗，传授犯罪方法，制作或者销售违禁物品、管制物品等违法犯罪活动的网站、通讯群组，不得利用网络发布涉及实施诈骗，制作或者销售违禁物品、管制物品以及其他违法犯罪活动的信息。

第四十九条　网络运营者应当加强对其用户发布的信息的管理，发现法律、行政法规禁止发布或者传输的信息的，应当立即停止传输该信息，采取消除等处置措施，防止信息扩散，保存有关记录，并向有关主管部门报告。

第五十条　任何个人和组织发送的电子信息、提供的应用软件，不得设置恶意程序，不得含有法律、行政法规禁止发布或者传输的信息。

电子信息发送服务提供者和应用软件下载服务提供者，应当履行安全管理义务，知道其用户有前款规定行为的，应当停止提供服务，采取消除等处置措施，保存有关记录，并向有关主管部门报告。

第五十一条　网络运营者应当建立网络信息安全投诉、举报制度，公布投诉、举报方式等信息，及时受理并处理有关网络信息安全的投诉和举报。

网络运营者对网信部门和有关部门依法实施的监督检查，应当予以配合。

第五十二条　国家网信部门和有关部门依法履行网络信息安全监督管理职责，发现法律、行政法规禁止发布或者传输的信息的，应当要求网络运营者停止传输，采取消除等处置措施，保存有关记录；对来源于中华人民共和国境外的上述信息，应当通知有关

机构采取技术措施和其他必要措施阻断传播。

第四章 监测预警与应急处置

第五十三条　国家建立网络安全监测预警和信息通报制度。国家网信部门应当统筹协调有关部门加强网络安全信息收集、分析和通报工作，按照规定统一发布网络安全监测预警信息。

第五十四条　负责关键信息基础设施安全保护工作的部门，应当建立健全本行业、本领域的网络安全监测预警和信息通报制度，并按照规定报送网络安全监测预警信息。

第五十五条　国家网信部门协调有关部门建立健全网络安全风险评估和应急工作机制，制定网络安全事件应急预案，并定期组织演练。

负责关键信息基础设施安全保护工作的部门应当制定本行业、本领域的网络安全事件应急预案，并定期组织演练。

网络安全事件应急预案应当按照事件发生后的危害程度、影响范围等因素对网络安全事件进行分级，并规定相应的应急处置措施。

第五十六条　网络安全事件发生的风险增大时，省级以上人民政府有关部门应当按照规定的权限和程序，并根据网络安全风险的特点和可能造成的危害，采取下列措施：

（一）要求有关部门、机构和人员及时收集、报告有关信息，加强对网络安全风险的监测；

（二）组织有关部门、机构和专业人员，对网络安全风险信息进行分析评估，预测

事件发生的可能性、影响范围和危害程度；

（三）向社会发布网络安全风险预警，发布避免、减轻危害的措施。

第五十七条　发生网络安全事件，应当立即启动网络安全事件应急预案，对网络安全事件进行调查和评估，要求网络运营者采取技术措施和其他必要措施，消除安全隐患，防止危害扩大，并及时向社会发布与公众有关的警示信息。

第五十八条　省级以上人民政府有关部门在履行网络安全监督管理职责中，发现网络存在较大安全风险或者发生安全事件的，可以按照规定的权限和程序对该网络的运营者的法定代表人或者主要负责人进行约谈。网络运营者应当按照要求采取措施，进行整改，消除隐患。

第五十九条　因网络安全事件，发生突发事件或者生产安全事故的，应当依照《中华人民共和国突发事件应对法》、《中华人民共和国安全生产法》等有关法律、行政法规的规定处置。

第六十条　因维护国家安全和社会公共秩序，处置重大突发社会安全事件的需要，经国务院决定或者批准，可以在特定区域对网络通信采取限制等临时措施。

第五章 法律责任

第六十一条　网络运营者不履行本法第二十三条、第二十七条规定的网络安全保护义务的，由有关主管部门责令改正，给予警告，可以处一万元以上五万元以下罚款；拒不改正或者导致危害网络安全等后果的，处五万元以上五十万元以下罚款，对直接负责的主管人员和其他直接责任人员处一万元以上十万元以下罚款。

关键信息基础设施的运营者不履行本法第三十五条、第三十六条、第三十八条、第四十条规定的网络安全保护义务的，由有关主管部门责令改正，给予警告，可以处五万元以上十万元以下罚款；拒不改正或者导致危害网络安全等后果的，处十万元以上一百万元以下罚款，对直接负责的主管人员和其他直接责任人员处一万元以上十万元以下罚款。

有前两款行为，造成大量数据泄露、关键信息基础设施丧失局部功能等严重危害网络安全后果的，由有关主管部门处五十万元以上二百万元以下罚款，对直接负责的主管人员和其他直接责任人员处五万元以上二十万元以下罚款；造成关键信息基础设施丧失主要功能等特别严重危害网络安全后果的，处二百万元以上一千万元以下罚款，对直接负责的主管人员和其他直接责任人员处二十万元以上一百万元以下罚款。

第六十二条　违反本法第二十四条第一款、第二款和第五十条第一款规定，有下列行为之一的，由有关主管部门责令改正，给予警告；拒不改正或者导致危害网络安全等后果的，处五万元以上五十万元以下罚款，对直接负责的主管人员处一万元以上十万元以下罚款：

（一）设置恶意程序的；

（二）对其产品、服务存在的安全缺陷、漏洞等风险未立即采取补救措施，或者未按照规定及时告知用户并向有关主管部门报告的；

（三）擅自终止为其产品、服务提供安全维护的。

有前款第一项、第二项行为，造成本法第六十一条第三款规定的后果的，依照该款规定处罚。

第六十三条　违反本法第二十五条规定，销售或者提供未经安全认证、安全检测或者安全认证不合格、安全检测不符合要求的网络关键设备和网络安全专用产品的，由有关主管部门责令停止销售或者提供，给予警告，没收违法所得；没有违法所得或者违法所得不足十万元的，并处二万元以上十万元以下罚款；违法所得十万元以上的，并处违法所得一倍以上五倍以下罚款；情节严重的，并可以责令暂停相关业务、停业整顿、吊销相关业务许可证或者吊销营业执照。法律、行政法规另有规定的，依照其规定。

第六十四条　网络运营者违反本法第二十六条第一款规定，未要求用户提供真实身份信息，或者对不提供真实身份信息的用户提供相关服务的，由有关主管部门责令改正；拒不改正或者情节严重的，处五万元以上五十万元以下罚款，并可以责令暂停相关业务、停业整顿、关闭网站或者应用程序、吊销相关业务许可证或者吊销营业执照，对直接负责的主管人员和其他直接责任人员处一万元以上十万元以下罚款。

第六十五条　违反本法第二十八条规定，开展网络安全认证、检测、风险评估等活动，或者向社会发布系统漏洞、计算机病毒、网络攻击、网络侵入等网络安全信息的，由有关主管部门责令改正，给予警告，可以处一万元以上十万元以下罚款；拒不改正或者情节严重的，处十万元以上一百万元以下罚款，并可以责令暂停相关业务、停业整顿、关闭网站或者应用程序、吊销相关业务许可证或者吊销营业执照，对直接负责的主管人员和其他直接责任人员处一万元以上十万元以下罚款。

有前款行为，造成本法第六十一条第三款规定的后果的，依照该款规定处罚。

第六十六条　违反本法第二十九条规定，从事危害网络安全的活动，或者提供专门用于从事危害网络安全活动的程序、工具，或者为他人从事危害网络安全的活动提供技术支持、广告推广、支付结算等帮助，尚不构成犯罪的，由公安机关没收违法所得，处五日以下拘留，可以并处五万元以上五十万元以下罚款；情节较重的，处五日以上十五日以下拘留，可以并处十万元以上一百万元以下罚款。

单位有前款行为的，由公安机关没收违法所得，处十万元以上一百万元以下罚款，并对直接负责的主管人员和其他直接责任人员依照前款规定处罚。

违反本法第二十九条规定，受到治安管理处罚的人员，五年内不得从事网络安全管理和网络运营关键岗位的工作；受到刑事处罚的人员，终身不得从事网络安全管理和网络运营关键岗位的工作。

第六十七条　关键信息基础设施的运营者违反本法第三十七条规定，使用未经安全审查或者安全审查未通过的网络产品或者服务的，由有关主管部门责令限期改正、停止使用、消除对国家安全的影响，处采购金额一倍以上十倍以下罚款，对直接负责的主管人员和其他直接责任人员处一万元以上十万元以下罚款。

第六十八条　违反本法第四十八条规定，设立用于实施违法犯罪活动的网站、通讯群组，或者利用网络发布涉及实施违法犯罪活动的信息，尚不构成犯罪的，由公安机关处五日以下拘留，可以并处一万元以上十万元以下罚款；情节较重的，处五日以上十五日以下拘留，可以并处五万元以上五十万元以下罚款。关闭用于实施违法犯罪活动的网站、通讯群组。

单位有前款行为的，由公安机关处十万元以上五十万元以下罚款，并对直接负责的主管人员和其他直接责任人员依照前款规定处罚。

第六十九条　网络运营者违反本法第四十九条规定，对法律、行政法规禁止发布或者传输的信息未停止传输、采取消除等处置措施、保存有关记录、向有关主管部门报告，或者违反本法第五十二条规定，不按照有关部门的要求对法律、行政法规禁止发布或者传输的信息停止传输、采取消除等处置措施、保存有关记录的，由有关主管部门责令改正，给予警告、予以通报，可以处五万元以上五十万元以下罚款；拒不改正或者情节严

重的，处五十万元以上二百万元以下罚款，并可以责令暂停相关业务、停业整顿、关闭网站或者应用程序、吊销相关业务许可证或者吊销营业执照，对直接负责的主管人员和其他直接责任人员处五万元以上二十万元以下罚款。

有前款行为，造成特别严重影响、特别严重后果的，由有关主管部门处二百万元以上一千万元以下罚款，责令暂停相关业务、停业整顿、关闭网站或者应用程序、吊销相关业务许可证或者吊销营业执照，对直接负责的主管人员和其他直接责任人员处二十万元以上一百万元以下罚款。

电子信息发送服务提供者、应用软件下载服务提供者，不履行本法第五十条第二款规定的安全管理义务的，依照前两款规定处罚。

第七十条　网络运营者违反本法规定，有下列行为之一的，由有关主管部门责令改正；拒不改正或者情节严重的，处五万元以上五十万元以下罚款，对直接负责的主管人员和其他直接责任人员，处一万元以上十万元以下罚款：

（一）拒绝、阻碍有关部门依法实施的监督检查的；

（二）拒不向公安机关、国家安全机关提供技术支持和协助的。

第七十一条　有下列行为之一的，依照有关法律、行政法规的规定处理、处罚：

（一）发布或者传输本法第十三条第二款和其他法律、行政法规禁止发布或者传输的信息的；

（二）违反本法第二十四条第三款、第四十三条至第四十五条规定，侵害个人信息权益的；

（三）违反本法第三十九条规定，关键信息基础设施的运营者在境外存储个人信息和重要数据，或者向境外提供个人信息和重要数据的。

违反本法第四十六条规定，窃取或者以其他非法方式获取、非法出售或者非法向他人提供个人信息，尚不构成犯罪的，由公安机关依照有关法律、行政法规的规定处罚。

第七十二条　有本法规定的违法行为的，依照有关法律、行政法规的规定记入信用档案，并予以公示。

第七十三条　违反本法规定，但具有《中华人民共和国行政处罚法》规定的从轻、减轻或者不予处罚情形的，依照其规定从轻、减轻或者不予处罚。

第七十四条　国家机关政务网络的运营者不履行本法规定的网络安全保护义务的，由其上级机关或者有关机关责令改正；对直接负责的主管人员和其他直接责任人员依法给予处分。

第七十五条　网信部门和有关部门违反本法第三十二条规定，将在履行网络安全保护职责中获取的信息用于其他用途的，对直接负责的主管人员和其他直接责任人员依法给予处分。

网信部门和有关部门的工作人员玩忽职守、滥用职权、徇私舞弊，尚不构成犯罪的，依法给予处分。

第七十六条　违反本法规定，给他人造成损害的，依法承担民事责任。

违反本法规定，构成违反治安管理行为的，依法给予治安管理处罚；构成犯罪的，

依法追究刑事责任。

第七十七条　境外的机构、组织、个人从事危害中华人民共和国网络安全的活动的，依法追究法律责任；造成严重后果的，国务院公安部门和有关部门并可以决定对该机构、组织、个人采取冻结财产或者其他必要的制裁措施。

第六章 附 则

第七十八条　本法下列用语的含义：

（一）网络，是指由计算机或者其他信息终端及相关设备组成的按照一定的规则和程序对信息进行收集、存储、传输、交换、处理的系统。

（二）网络安全，是指通过采取必要措施，防范对网络的攻击、侵入、干扰、破坏和非法使用以及意外事故，使网络处于稳定可靠运行的状态，以及保障网络数据的完整性、保密性、可用性的能力。

（三）网络运营者，是指网络的所有者、管理者和网络服务提供者。

（四）网络数据，是指通过网络收集、存储、传输、处理和产生的各种电子数据。

（五）个人信息，是指以电子或者其他方式记录的能够单独或者与其他信息结合识别自然人个人身份的各种信息，包括但不限于自然人的姓名、出生日期、身份证件号码、个人生物识别信息、住址、电话号码等。

第七十九条　存储、处理涉及国家秘密信息的网络的运行安全保护，除应当遵守本法外，还应当遵守保密法律、行政法规的规定。

第八十条　军事网络的安全保护，由中央军事委员会另行规定。

第八十一条　本法自 2017 年 6 月 1 日起施行。

《数据出境安全管理政策问答（2025 年 10 月）》

发文机构：国家互联网信息办公室

发布时间：2025.10.31

生效时间：2025.10.31

国家互联网信息办公室持续加强数据出境安全管理政策宣贯，指导和帮助数据处理者高效合规开展数据出境活动。经对近期收到的咨询问题进行研究，现将一些有代表性的问题和答复公布如下。

1.《促进和规范数据跨境流动规定》明确了“跨境购物、跨境寄递、.....、考试服务等”豁免场景，其中“等”字应如何准确理解？

答：《促进和规范数据跨境流动规定》第五条第一款第（一）项规定，“为订立、履行个人作为一方当事人的合同，如跨境购物、跨境寄递、跨境汇款、跨境支付、跨境开户、机票酒店预订、签证办理、考试服务等，确需向境外提供个人信息的可免予安全评估、订立合同或认证”，此处“等”字理解为可以纳入豁免情形的，不限于以上所列情形。

但需注意的是，豁免场景必须同时满足两个条件：

1）系为订立、履行个人作为一方当事人的合同；

2）确需向境外提供个人信息，此处可根据有关法律法规、规章、规范性文件、国家标准和实际出境场景等对最小必要个人信息范围进行判断。

同时，根据《促进和规范数据跨境流动规定》第十条，数据处理者向境外提供个人信息的，应当按照法律、行政法规的规定履行告知、取得个人单独同意、进行个人信息保护影响评估等义务。

2.酒店行业中“境内个人预定境内酒店”的数据是否符合《促进和规范数据跨境流动规定》第五条豁免场景？

答：根据《促进和规范数据跨境流动规定》第五条，酒店企业在境内个人预定境内酒店时出境个人信息不符合“为订立、履行个人作为一方当事人的合同，确需向境外提供个人信息”的情形，不免予申报数据出境安全评估、订立个人信息出境标准合同、通过个人信息保护认证。

3.向境外提供员工的身份证、护照和银行账户是否适用于豁免规定“按照依法制定的劳动规章制度和依法签订的集体合同实施跨境人力资源管理，确需向境外提供员工个人信息的”范畴？

答：根据《促进和规范数据跨境流动规定》第五条第一款第（二）项，“按照依法制定的劳动规章制度和依法签订的集体合同实施跨境人力资源管理，确需向境外提供员工个人信息”的情形免予申报数据出境安全评估、订立个人信息出境标准合同、通过个人信息保护认证。数据处理者为实施人力资源管理所必需处理个人信息，要按照依法制定的劳动规章制度和依法签订的集体合同实施。劳动规章制度和集体合同中的相关规定和约定应当遵守个人信息处理的原则和规则，特别是要符合必要、目的明确、最小化处理等原则，只能处理与实施人力资源管理目的直接相关的个人信息，采取对个人权益影响最小的方式。身份证、护照、银行账户等员工个人信息是否属于“确需”范围，应从上述角度进行具体判断。

4.数据处理者在被告知有“重要数据”后应当在两个月内通过所在地省级网信部门向国家网信部门申报数据出境安全评估，是否可以提供更多时间和灵活性？

答：数据处理者被告知掌握重要数据或者掌握的数据被公开发布为重要数据后，如需继续开展相关数据出境活动的，应当在被告知或者公开发布后 2 个月内，通过所在地省级网信部门向国家网信部门申报数据出境安全评估。如出境场景复杂度较高、所需准备评估材料时间较长，建议数据处理者在重要数据识别认定期间，同步梳理业务场景，并准

备相关申报材料，在重要数据被认定告知后，抓紧按程序申报数据出境安全评估。

5.《数据出境安全评估申报指南（第三版）》中“数据处理者收集和产生的数据存储在境内，境外的机构、组织或者个人可以查询、调取、下载、导出”的“境外”是指什么？

答：“数据处理者收集和产生的数据存储在境内，境外的机构、组织或者个人可以查询、调取、下载、导出”的“境外”是指数据访问或调用行为发生在境外。境外机构、组织的工作人员在境内查询、调取、下载、导出调用数据处理者存储在境内的数据、没有将数据传输至境外，不属于数据出境活动。

6.数据出境场景、接收方均不变，但系统可能会升级或者更换，数据处理者是否需要重新申报数据出境安全评估？

答：根据《数据出境安全评估办法》第十四条规定，在有效期内出现以下情形之一的，数据处理者应当重新申报评估：（一）向境外提供数据的目的、方式、范围、种类和境外接收方处理数据的用途、方式发生变化影响出境数据安全的，或者延长个人信息和重要数据境外保存期限的；（二）境外接收方所在国家或者地区数据安全保护政策法规和网络安全环境发生变化以及发生其他不可抗力情形、数据处理者或者境外接收方实际控制权发生变化、数据处理者与境外接收方法律文件变更等影响出境数据安全的；（三）出现影响出境数据安全的其他情形。

如数据处理者调整数据出境相关系统，需按照相关条款规定进行判断，如未出现上述情形，则无需重新申报数据出境安全评估。

7.个人信息处理者开展的业务活动涉及持续出境个人信息的情况，需要多次进行个人信息出境标准合同备案吗？

答：个人信息处理者开展的业务活动如仅涉及同一境外接收方，且预计每年出境个人信

息的数量符合订立标准合同的规定条件，可在合理预测出境个人信息数量基础上备案一次合同。如果自当年 1 月 1 日起累计出境个人信息数量达到申报数据出境安全评估的规定条件，个人信息处理者应当通过所在地省级网信办向国家网信办申报安全评估。

8.个人信息处理者完成个人信息出境标准合同备案后，在什么情形下需要重新订立标准合同，完成备案后因业务发展需要新增少量个人信息出境场景，该如何处理？

答：《个人信息出境标准合同办法》第八条规定，“在标准合同有效期内出现下列情形之一的，个人信息处理者应当重新开展个人信息保护影响评估，补充或者重新订立标准合同，并履行相应备案手续：（一）向境外提供个人信息的目的、范围、种类、敏感程度、方式、保存地点或者境外接收方处理个人信息的用途、方式发生变化，或者延长个人信息境外保存期限的；（二）境外接收方所在国家或者地区的个人信息保护政策和法规发生变化等可能影响个人信息权益的；（三）可能影响个人信息权益的其他情形”。完成备案后，若新增个人信息出境场景符合上述情形之一，则应当重新开展个人信息保护影响评估，补充或者重新订立标准合同并履行备案义务。

9.境外接收方可以将个人信息处理者通过订立个人信息出境标准合同方式出境的个人信息再提供给境外第三方吗？

答：如境外接收方需将个人信息处理者通过订立个人信息出境标准合同方式出境的个人信息提供给境外第三方，个人信息处理者和境外接收方在订立个人信息出境标准合同时，应在个人信息出境标准合同范本的附录一《个人信息出境说明》“（六）境外接收方只向以下中华人民共和国境外第三方提供个人信息（如适用）”部分予以说明。

10.《个人信息出境认证办法》施行后，相关认证参照的依据和标准主要是？

答：认证机构开展个人信息出境认证、相关企业申请个人信息出境认证，参照的依据和

标准主要是 2022 年 11 月公布的《关于实施个人信息保护认证的公告》以及国家标准《数据安全技术 个人信息跨境处理活动安全认证要求》（GB/T 46068-2025）。此外，落实《个人信息出境认证办法》规定，国家互联网信息办公室将按程序公示相关专业认证机构，具体公示信息请及时关注中国网信网。

二、实务解读

1. 企业在中国境内部署及应用 AI Agent 的主要法律问题（一）

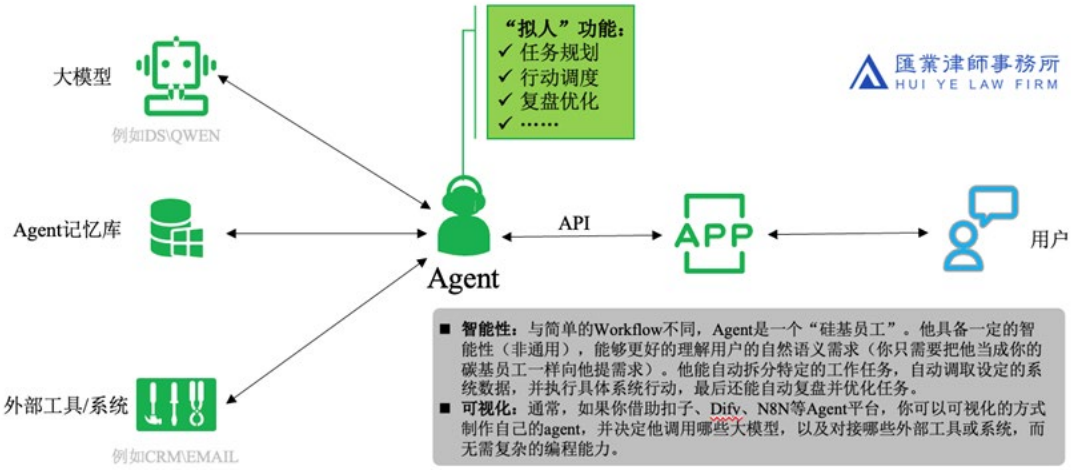
供稿人：黄春林（上海市汇业律师事务所）、袁姜涛（上海市汇业律师事务所）、钱静雯（上海市汇业律师事务所）

随着企业降本增效需求的不断增强，以及供给端 AI 能力成熟度不断提高， 2026 年注定将成为企业部署及应用 AI 的爆发之年。其中，具备自主决策与跨系统协同能力的 AI Agent，正加速进入企业实际业务场景，将在生产力层面推动企业真正开始从“数字化转型”到“智能化转型”。然而，关于 Agent 的法律定性、监管框架，以及 Agent 带来的算法安全、责任边界、竞争合规、用户权利保障及数据合规等问题，也会给企业带来较大的法律挑战。

参考中国近期立法及监管实践，结合近期类似项目经验及行业实践，我们简要分析企业在中国境内应用 AI Agent 的主要法律问题如下，仅供参考。

一、Agent 的主要结构、功能及实例

简化理解，Agent 就相当于企业的一个“硅基员工”，那些本来应当招聘一个员工干活的场景（通常是那些“低智流程化的场景”，例如客服、仓管等场景），可以通过部署一个 Agent 来替代。为方便理解，我们将 Agent 的主要结构及功能简化为如下图示：



以部署 Agent 的智能客服为例：当消费者通过品牌 APP 给智能客服提问“为什么我的积分还没到账？”时，APP 通过 API 调用的 Agent 就开始如下操作了：

（1）自行或调用大模型理解“为什么我的积分还没到账？”这段自然语义是什么意思，即对应什么任务？

（2）看看这个任务在 Agent 记忆库里能不能找到类似答案（像法务拿到消费者投诉后，去查询下以前是否有同样的问题咨询过外部律师一样……）

（3）将这个任务拆解（像法务牵头分拆任务，分别给各部门派活儿），例如会员身份查询、订单记录查询、物流记录查询、积分规则查询、文字答复消费者等；

（4）然后分别从各个系统（例如 CRM 系统、POS 系统、OMS 系统等）查询数据并汇总判断（就像法务分别从电商、客服、物流等团队要数据），生成初步结论；

（5）再把初步结论丢给大模型生成消费者友好的语言（就像法务把初稿给外部律师一样）；

（6）再拉起邮件接口或者客服系统接口，回复给消费者；

（7）还没解决的话，继续以上步骤，周而复始，或者拉起人工客服；等等。

二、Agent 的法律定性及主要法律风险

从以上 Agent 结构、功能不难理解， Agent 不是一个单一的智能工具，而是由大模型、任务规划模块、工具调用模块、任务执行模块、长期/短期记忆模块等组成的复杂智能系统，具备“理解—规划—执行—反馈”的智能闭环能力。因此，结合其不同结构与功能等因素，根据我国当前立法及监管实践，我们简要分析 Agent 不同模块的核心法律属性及主要风险如下：

模块	核心法律定性	主要法律风险
大模型 （外接或嵌入）	✓ 计算机程序 ✓ 信息服务 ✓ 生成式人工智能	☐ 算法安全 ☐ 伦理合规 ☐ 信息安全 ☐ 数据合规 ☐ 侵权/违约责任 ☐ 政府合规
任务规划模块	✓ 计算机程序 ✓ 自动化决策	☐ 算法安全（确定性及蔓延风险等） ☐ 伦理合规 ☐ 用户知情与决定权 ☐ 政府合规
工具调用模块	✓ 计算机程序接口	☐ 网络安全 ☐ 侵权/违约责任（互操作性风险等）
任务执行模块	✓ 计算机程序 ✓ 信息服务 ✓ 生成式人工智能	☐ 网络安全 ☐ 信息安全 ☐ 用户权益
长期/短期记忆模块	✓ 数据 ✓ 信息 ✓ 数据中心	☐ 数据合规 ☐ 信息安全 ☐ 政府合规

当然，以上仅仅是基于典型 Agent 的抽象理解，实际上在具体的业务场景中，每个 Agent 受限于不同的技术框架、生态链、落地场景、模态类型及数据流等影响，可能会产生新的法律定性并暴露新的法律风险。

三、企业内部应用 Agent 的主要合规义务

根据《网络安全法》《数据安全法》《个人信息保护法》《互联网信息服务算法推荐管理规定》《互联网信息服务深度合成管理规定》《生成式人工智能服务管理暂

行办法》等规定，我们梳理的企业内部部署并应用 Agent 的主要合规义务包括但不限于：

- 应当对 Agent 开展算法安全评估及伦理审查，确保算法机制机理安全，建立算法确定性护栏和推理式防御，建立 Agent 身份核验机制，有效应对 Agent 安全蔓延，依法开展科技伦理审查；
- 采取多 Agent 协同架构的，应当通过规程、协议或技术护栏等方式明确每个 Agent（或 Agent 供应商）的权限及责任；
- 确保调用的第三方大模型、使用的外部 Agent 在中国境内具有运营合法性，且商用大模型具有合法的授权，或者遵守开源大模型的开源协议；
- 通过第三方平台部署 Agent 的，应当依法审查 Agent 平台的电信业务等资质和网络安全能力；
- 确保调用的第三方接口或 MCP 具有合法的授权，相关的互操作遵守适用的开放平台协议、规则或政策；
- 与外部工具或系统之间存在数据交互时，避免发生监管数据（个人信息或重要数据）跨境传输，除非依法履行了必要的合规手续；
- 确保大模型、Agent、RAG 相关的数据具有合法来源，数据处理方式符合法律规定；
- 如涉及处理外部用户或内部员工的个人信息，应当依法履行告知/同意合规责任，依法开展 PIA；
- 落实用户实名制要求，依法留存相关网络日志；
- 加强信息安全管理，建立健全用于识别违法和不良信息的特征库，及时处置违法违规内容；
- 应当建立适当的人工干预/审查机制，确保对外提供的 Agent 工作成果符合法律规定；
- 建立健全算法安全相关公司管理制度，设置与业务相适应的岗位及人员；等。

2. 企业在中国境内部署及应用 AI Agent 的主要法律问题（二）

供稿人：黄春林（上海市汇业律师事务所）、袁姜涛（上海市汇业律师事务所）、钱静雯（上海市汇业律师事务所）

随着企业降本增效需求的不断增强，以及供给端 AI 能力成熟度不断提高，2026 年注定将成为企业部署及应用 AI 的爆发之年。其中，具备自主决策与跨系统协同能力的 AI Agent，正加速进入企业实际业务场景，将在生产力层面推动企业真正开始从“数字化转型”到“智能化转型”。然而，关于 Agent 的法律定性、监管框架，以及 Agent 带来的算法安全、责任边界、竞争合规、用户权利保障及数据合规等问题，也会给企业带来较大的法律挑战。

参考中国近期立法及监管实践，结合近期类似项目经验及行业实践，我们简要分析企业在中国境内应用 AI Agent 的主要法律问题如下，仅供参考。

一、企业应用 Agent 对外提供服务的主要合规义务

企业部署并应用 Agent 对外提供服务的（例如前面的智能客服实例），除了应当遵守第三部分的合规义务外，还应当遵守的主要合规义务包括但不限于：

1. 算法备案/合成生成备案

根据《互联网信息服务算法推荐管理规定》《互联网信息服务深度合成管理规定》等规定，Agent 具有舆论属性或者社会动员能力的（实践中把握的口径很宽泛），应当依法办理算法备案。

实践中，根据 Agent 的功能不一样，我们建议针对性办理如下类别备案：

主要功能描述	备案类型建议	行业实践
自动化调度决策类 Agent	互联网信息服务算法备案	➢ 华之道智能体调度算法，网信算备 440112585961406250013 号； ➢ 新灵犀 AI 智能体全资源调度决策算法，网信算备 110108990813806250015 号
深度合成类 Agent（服务提供者）	深度合成服务算法备案	➢ 爱原生 AI 端侧 Agent 内容生成算法，网信算备 440309232851101250019 号； ➢ 蚁群 AI 智能体文本生成算法，网信算备 440106089690101250015 号； ➢ 智识神工 CAMo 编程智能体模型生成算法，网信算备 310115727610601250015 号； ➢ Monica 智能对话生成算法，网信算备 110108429778801250019 号
深度合成类 Agent（技术支持者）	深度合成服务算法备案	➢ TSAgent 大模型生成合成算法，网信算备 440115168425601250017 号； ➢ 星睿多智能体协作内容生成算法，网信算备 110112803015101250013 号

*实践中，若 Agent 部署在境外的，可能无法完成备案。

2. 生成式人工智能备案/登记

根据《互联网信息服务算法推荐管理规定》《生成式人工智能服务管理暂行办法》《具有舆论属性或社会动员能力的互联网信息服务安全评估规定》等规定，Agent 具有舆论属性或者社会动员能力，且内嵌或调用了第三方大模型的，应当依法办理生成式人工智能备案/登记。

实践中，根据大模型的基础合规情况，我们建议针对性办理如下备案/登记：

大模型情况	类型建议	行业实践
调用第三方大模型已经完成备案（无微调等）	成式人工智能登记	➢ 携程AI行程助手，Shanghai-XingChengZhuShou-20250714S0096； ➢ 最美AI智能体，Shanghai-ZuiMeiAIZhiNengTi-20250911S0109； ➢ 万兴超媒Agent，Hunan-wsSuperMedia-20250905S0005
调用第三方大模型已经完成备案（有微调等）	成式人工智能备案	➢ 特斯拉xBot客户服务，Shanghai-TeslaxBotKeHuFuWu-202507230067； ➢ 奔驰虚拟助手，Beijing-BenChiXuNiZhuShou-202509080121； ➢ 安利AI助手，Guangdong-amway-202509120096
内嵌或调用第三方大模型未完成备案	成式人工智能备案	年初火极一时的 Manus，据传因为调用境外大模型 Claude 而无法完成大模型备案（目前仍未查询到其备案记录），因此无法在中国境内提供服务；相反的，该公司运营的 Monica 因为调用了境内合规的 DeepSeek，因此完成了算法备案。

*实践中，若 Agent 部署在境外，或者调用境外的大模型的，可能无法完成备案/登记。

3. 其他合规义务

- 应当由法务、合规人员参与 Agent 设计，充分发挥专业合规经验，利用 Agent 的任务边界、安全围栏、限制语、拒答策略等控制业务流确定性，事前规避潜在法律风险，确保 Agent 不越权、不越界、不误导等。
- 应当建立越狱话术及提示词注入攻击的识别与防范机制，完善意图对齐及约束遵循机制，确保信息内容安全，防止敏感数据及商业秘密泄露；
- Agent 上线前，应当参照《网络安全技术 生成式人工智能服务安全基本要求》等要求开展安全评估；
- 应当在宿主 APP/小程序等的显著位置标明算法备案编号；

- 以算法规则、用户协议等方式公示算法服务的基本原理、目的意图和主要运行机

制；

- 在处理用户个人信息前，必须清晰、明确地告知处理其个人信息的目的、方式和范围（尤其是否用于模型训练、Agent 记忆库等），并取得用户的有效同意；
- 通过 API 接口或 MCP 向第三方提供个人信息的，应当在共享清单中依法列明；

- 应当向用户提供不针对其个人特征的选项，或者向用户提供便捷的关闭算法推荐服务（若有）的选项；
- 应当保护消费者公平交易的权利，不得根据消费者的偏好、交易习惯等特征，利用 Agent 在交易价格等交易条件上实施不合理的差别待遇等违法行为；
- 在 Agent 最终生成合成内容的文件元数据中添加隐式标识，并在特定情形下对生成内容添加显式标识，避免公众混淆或误认；
- 建立健全投诉、举报机制，设置便捷的投诉、举报入口；等。

二、企业通过第三方平台部署 Agent 的主要法律问题

1. 国内主流的 Agent 平台

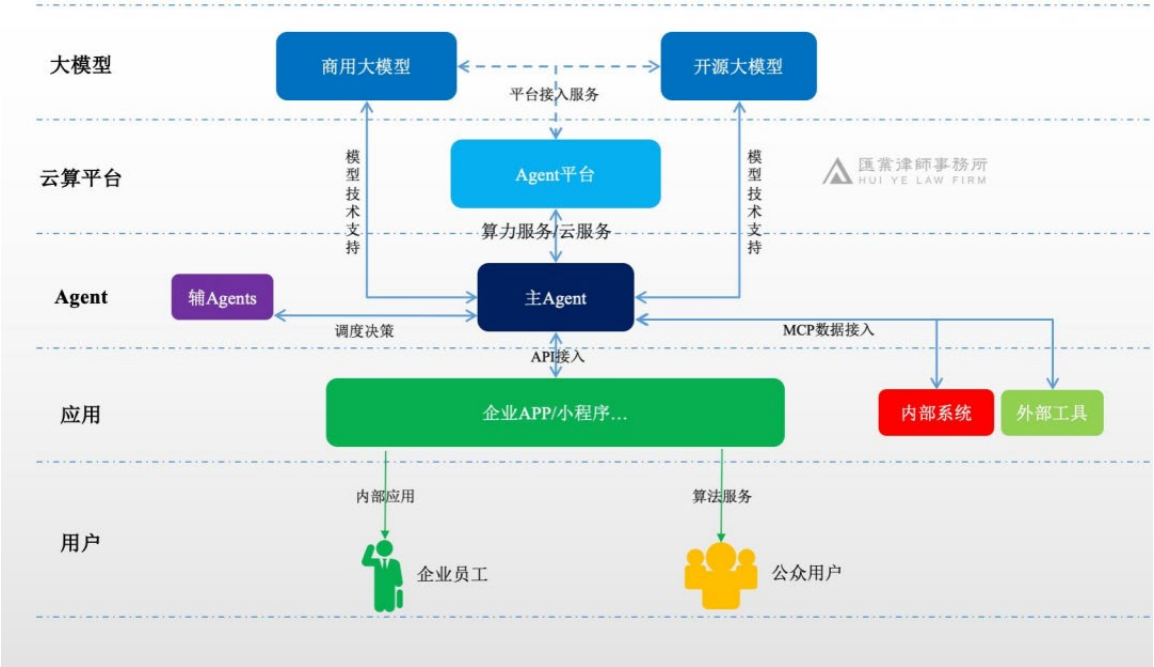
由于平台通常提供成熟的大模型介入、Agent 编排框架、运维及安全支持、弹性运算能力等，因此企业通过第三方平台部署 Agent，可显著降低技术与运维门槛。目前，中国市场上主流的 Agent 平台包括但不限于：

Agent 平台名称	法律实体	协议
阿里云百炼	通义云启（杭州）信息技术有限公司	大模型服务平台百炼—服务协议
百度千帆	百度在线网络技术（北京）有限公司	大模型服务及 Agent 开发平台用户协议 大模型服务及 Agent 开发平台专项约定
字节扣子	北京春田知韵科技有限公司 北京火山引擎科技有限公司	扣子用户协议、扣子订阅版专用条款

2. Agent 平台的法律地位

参照目前主流 Agent 平台法律框架，通过第三方平台部署 Agent 的架构示意图如

下：



从上图不难看出，自有 Agent 通过宿主 APP 对外实际提供算法服务。因此，与算法相关的合规义务（包括但不限于算法备案/登记、内容安全管理及风险防控等）应由 APP 运营主体承担；参照当前监管实践，Agent 由第三方供应商提供的，供应商可能需要依法办理技术支持类合成生成算法备案。

在该模式下，Agent 平台主要为企业部署 Agent 及调用大模型提供算力、云资源等基础设施层面的支持服务（另有明确提供其他服务的情形除外），并不直接以大模型服务提供者或算法技术支持者的身份对外提供服务。对此，各主要平台通常已在其用户协议及平台规则中予以反复明确和声明。

3. 主要法律风险点

企业通过第三方平台部署 Agent，虽然在技术选型和部署模式上具备更高的灵活性，但相应也会暴露出更多潜在的安全与法律风险。基于风险可控原则，具备条件的企业在 Agent 发展的当前阶段，宜优先考虑采用私有化部署模式（或者至少要在最终

输出环节增加额外的安全模型）。主要风险包括但不限于：

(1) 商业秘密泄露风险：Agent 的编排逻辑、提示词工程（Prompt Engineering）及业务规则配置，可能涉及企业核心业务逻辑与商业秘密，在第三方平台环境下存在被不当获取或使用的风险。

(2) 内部系统接口暴露风险：Agent 对企业内部 CRM、POS 等系统接口的调用，可能扩大攻击面，进而引发数据泄露、数据篡改、网络攻击、黑灰产利用或电信诈骗等安全风险。

(3) 外部工具链路不透明风险：Agent 调用的外部工具或 MCP 服务链路不透明、服务安全性不足，可能导致违法违规处理个人信息、数据安全事件或跨境数据合规风险。

(4) 数据处理授权偏离真实意思表示的风险：第三方 Agent 平台可能通过线上用户协议、平台规则或默认设置等方式，取得对 Agent 记忆库、RAG 数据、用户输入输出数据等的处理“形式授权”，存在与企业真实意思表示不一致的合规隐患。

(5) 平台锁定与迁移受限风险：部分 Agent 平台可能通过技术或协议安排限制 Agent 及相关数据的迁移、重构，影响企业的业务连续性和议价能力。

(6) 平台主体资质不足风险：个别 Agent 平台的实际运营主体或签约主体，可能未取得与其业务模式相匹配的必要资质或许可（如 IDC、ICP 等增值电信业务许可），从而给企业合作带来合规连带风险。

(7) 平台合规“连坐”与业务中断风险：若第三方 Agent 平台因自身合规问题被监管处罚、下架或整改，企业基于该平台运行的 Agent 可能被迫中止，影响核心业务连续性，且企业通常缺乏有效的替代方案或过渡机制。

(8) 平台未提供有效的合规能力风险：Agent 平台初期往往仅聚焦 Agent 业务编排能力，未完整提供企业落地法律合规责任的能力（例如生成内容标注、敏感权限管理、用户告知同意等），但承担监管压力与商业后果的最终可能是企业；等等。

六、法务合规在企业部署/应用 Agent 中的价值

Agent 本身的可视化及自然语义执行能力，大大降低了法务合规人员参与 Agent 部署和应用的技术门槛。因此，法务合规部门可以深度介入 Agent 的设计、训练、部署与运营过程，从“后置擦屁股”转向“前置划边界”，实现 Agent 全生命周期合规治理，确保业务创新在可控边界内运行，有效平衡业务效率和合规性。

法务参与 Agent 设计的优势体现在多个方面，包括但不限于：

(1) 熟悉法律红线与监管要求：能够协助明确 Agent 的职业边界和业务边界、拒答策略及约束遵循规则，擅长意图对齐并识别越狱话术，确保系统运行始终在合规框架内。

(2) 具备优秀的语义与逻辑能力：可深度参与上下文工程与提示词工程，制定限制语和安全围栏，编制生成内容测试题库、拒答/非拒答测试题库，从源头提升输出的可控性和业务确定性。

(3)**拥有丰富的风险识别与处置经验：**熟悉业务部门常见的风险点（毕竟经常干擦屁股的事儿），可通过安全模型训练、流程编排优化、提示词注入攻击预防及 RAG 等方式，提前规避 Agent 成为业务员工之外的“第二个草台班子”。

(4)**理解司法裁量与责任边界：**能够运用免责声明、勤勉合规设计等手段，帮助组织合理分散、降低因 Agent 翻车事件而引发的法律责任。

(5)**擅长结构化业务与合同设计：**能清晰界定不同模块、不同主体的权责边界，有利于构建透明、可追责的 Agent 体系。

(6)**擅长权利保护：**通过著作权、商业秘密及专利等适当的路径，有效保护企业的 Agent、Agent 记忆库、prompt 等合法权益，维持竞争优势。

(7)**具备应急响应能力：**在 Agent“放飞自我”时，能够迅速协调资源介入，必要时应对潜在的监管检查，降低事件扩散风险。

当然，以上这些优势也正是法务合规人员应当努力的方向。当前，Agent 模式方兴未艾，大家都在一条起跑线上。

综上，虽然 Agent 技术能够显著提升企业的效率与成本效益，但其并未从根本上解

决 AI 的不确定性。在当前技术成熟度下，一旦在生产环境中真实部署，可能引入更多潜在法律风险。**因此，企业上线 Agent 时，无论是构建“事前合规护栏”，还是承担“事后风险救火”，法务合规团队的专业价值始终不可或缺。**