



云计算标准和开源推进委员会
Open source and Standard for Cloud Advance Reform



数字政府建设赋能计划

政务大模型落地实践指南

云计算标准和开源推进委员会
数字政府建设赋能计划
2025年6月

编制说明

本指南的撰写得到了数字政府领域多家单位与专家的支持和帮助，主要参与单位如下：

中国信息通信研究院云计算与大数据研究所、宜兴市数据局、克拉玛依市数据资源和政务服务中心、广州市白云区城市管理和综合执法局、深圳市福田区政务服务和数据管理局。

主要参与企业如下：

华为云计算技术有限公司、宜兴市大数据发展有限公司、数字丝路新疆产业投资集团有限公司、贵州贵安发展集团有限公司、南京大数据集团有限公司、北京百度网讯科技有限公司、中电信数政科技有限公司、天翼云科技有限公司、浪潮云信息技术股份公司、中经网数据有限公司、湖北省楚天云有限公司、上海梦创双杨数据科技股份有限公司、蜜度科技股份有限公司、北京电子数智科技有限责任公司、中国联合网络通信有限公司智能城市研究院、浪潮软件集团有限公司。

主要参与专家如下：

栗蔚、徐恩庆、张琳琳、宋光通、吴佳兴、吴宁、宋佳明、王宁、李志强、仇俊、崔昊、孙腾中、刘灵娟、朱一飞、陈伟、张宝玉、白亮、滕希成、阎丰、傅鹏、邱泳钦、张英博、张亮、高红、刘增志、祁超、王鲁、苗子聪、孟子杰、李兆丽、张敏、唐晓东、于希光、戴鸿轶、朱璐、韩同、荆潇、伊丽娜、韩旭、柯鑫、王昉、宋汝良、何宁宁、郭真、王鹏、申奇、郭大字、陈兆亮、陈鉴、宁方刚。

前言

在新时代的征途上，我国正全面推进深化改革，力图在全球化浪潮中把握新的发展机遇。智能化作为推动经济社会发展的重要引擎，已被提升至国家战略层面。政务领域对于大模型等人工智能领域的研发与应用，正是顺应了数字经济高质量发展的时代要求，展现了政府在创新应用方面的前瞻性和行动力。大模型等创新技术应用于政务领域将致力于提高政务服务的效率和质量，并在多场景下展现出强大的应用潜力，成为数字化转型中的新趋势和亮点。随着《关于进一步优化政务服务提升行政效能 推动“高效办成一件事”的指导意见》、《关于推动未来产业创新发展的实施意见》、《生成式人工智能服务管理暂行办法》等一系列政策文件的密集出台，我国对政务的智能化发展提出了明确要求，旨在通过政策引导和技术赋能，加速传统产业的数字化转型，促进政务服务的智慧化升级。

在此背景下，云计算标准和开源推进委员会联合相关主管单位、科研机构 and 行业企业，依托各方在政务大模型领域的研究成果与实践经验，围绕政务大模型发展背景、政务大模型建设路径以及政务大模型未来发展趋势形成本报告，旨在为政务行业的智能化升级提供战略指引和实践参考，推动我国政务行业在智能化道路上趋深行稳。

目 录

一、政策产业齐驱，政务大模型场景应用“汇点成面”	1
（一）多项政策支持政务大模型场景化落地	1
（二）多地共同探索政务大模型试点性验证	3
（三）产业渗透度不断加深，创新突破频现	5
二、技术流程并重，政务大模型落地建设“规致有状”	8
（一）需求端：场景挖掘上，强调多部门协同共致	8
（二）供给端：技术服务上，强调全过程闭环落地	14
（三）生态端：行业生态上，强调标准化以评促优	26
三、四维要素共塑，政务大模型未来发展“趋势向善”	28
（一）构建高质量政务数据基础，增强政务大模型服务精准性 ..	28
（二）智能体驱动治理体系重构，强化政务大模型决策灵活性 ..	29
（三）筑牢全链路内生安全体系，提升政务大模型可靠可信性 ..	30
（四）创新跨部门组织协同方式，强化政务大模型服务敏捷性 ..	31

图 目 录

图 1 各省市部署 Deepseek 的部门数量分布图 (2025.2~4)	5
图 2 政务大模型先锋实践场景图	6
图 3 DeepSeek 政务部署应用类型统计图	7
图 4 政务大模型项目实施协作关系图 (需求侧)	13
图 5 政务大模型落地服务关键步骤图 (供给侧)	14
图 6 政务大模型标准体系 2.0 全景图	27

表 目 录

表 1 北京 2025 年 2 月~4 月 DeepSeek 应用案例	4
表 2 政务大模型落地实践重点工作表	8

一、政策产业齐驱，政务大模型场景应用“汇点成面”

（一）多项政策支持政务大模型场景化落地

近年来，我国陆续出台一系列指导意见及通知，持续加强对人工智能产业发展和技术创新的统筹指导，规范与推动人工智能应用建设，以实现 AI 与实体经济的深度融合。2024 年 5 月，国家数据局等四部委联合印发《关于深化智慧城市发展 推进城市全域数字化转型的指导意见》，鼓励发展基于人工智能等技术的智能分析、智能调度、智能监管、辅助决策，全面支撑赋能城市数字化转型场景建设与发展。2024 和 2025 年政府工作报告连续强调要推进“人工智能+”行动，与我国新质生产力培育战略形成深度耦合。2025 年 4 月，习近平总书记在上海考察时指出：“人工智能是年轻的事业，也是年轻人的事业。”我国正以人工智能等前沿技术突破为支点、制度创新为保障、青年人才为先锋，加速构建开放协同的人工智能创新生态，推动技术攻关与应用场景双向奔赴，为高质量发展锻造关键引擎，在全球科技革命浪潮中开辟中国人工智能发展的新境界。

同时，各地因地制宜出台一系列政策立体化助推大模型在数字政府领域场景化应用。这些政策以产业升级为核心驱动力，一方面引导传统产业智能化、数字化转型，培育新兴数字产业集群，激发地方经济新动能，增强区域经济竞争力；另一方面以技术赋能政务服务，重塑公共服务流程，打造高效、便捷、智能的新型政务服务体系，实现经济发展与民生福祉的协同提升。

2023年7月，北京市政务服务管理局在全球数字经济大会人工智能高峰论坛上，率先发布政务服务大模型场景需求清单，标志着地方政务大模型应用探索迈入新阶段。

2024年5月，《上海市推进“人工智能+”行动 打造“智慧好办”政务服务实施方案》正式发布，提出深化人工智能辅助申报，推动人工智能赋能政务服务需求侧，为企业群众提供智能预填、智能预审等智慧化服务。

2024年5月，《广东省关于人工智能赋能千行百业的若干措施》提出惠企利民建设智慧政府。利用政务大模型智能化升级广东政务服务网、“粤系列”政务服务平台，提供全时在线问答和搜索服务。

2024年7月，《北京市推动“人工智能+”行动计划（2024-2025年）》面向政务咨询、业务办理等场景，推进大模型技术在接诉即办智能受理、智能办理中的试点应用，准确、及时、便利回应群众办事诉求。做好政务办公助手，依托京智、京办、京通平台，接入政策问答、流程管理等领域的大模型工具。立足公共服务海量基础数据，构建开放式城市大模型服务平台，打造智慧城市大脑，提升公共服务效能。

2025年5月，山东省人民政府办公厅印发了《关于加快人工智能赋能重点领域高质量发展的推进方案》（以下简称《推进方案》）和《关于支持人工智能全产业链创新发展的若干政策措施》（以下简称《政策措施》）。其中《推进方案》提出打造智能高效便捷政务服务体系。打造政务服务“智能客服”应用场景，提升咨询、查询等服

务效率和智能化便捷化水平。打造 12345 热线智能办理助手，集成机器人应答、辅助填单、智能分类、智能转派等功能，提高市民诉求解决效率。打造“鲁惠通”政策精准推送应用场景。打造“数字机关”智能助手，建设公文智能写作、政策文件智能检索等共性功能，提高机关办公效率。《政策措施》中提出支持政务领域应用。统筹“数字山东”发展资金，支持省级利用“开源+闭源”大模型，集约建设“人工智能中台”及多模态数据集，构建一体化基础大模型服务体系，加强规范化、便利化人工智能支撑能力供给。

（二）多地共同探索政务大模型试点性验证

在 2025 年春节假期，国产人工智能大模型 DeepSeek 惊艳亮相，迅速引发社会广泛关注，热度持续攀升。该模型凭借强大的技术实力，正助力政务大模型领域掀起新一轮的升级热潮，推动数字政府向数智化、高效化方向迈进。政务领域在大模型选型策略上呈现出显著的异构化特征，需要针对垂直行业场景的差异化需求，遴选与领域特性相适配的模型能力，进行个性化部署。

以 DeepSeek 为例，根据公开信息整理，自 2025 年 2 月至 4 月上旬，DeepSeek 已覆盖全国 31 个省级行政区（含直辖市），累计部署超 190 余项场景，至少 30 余项部署明确采用“满血版”DeepSeek 模型，涵盖政务服务、政务办公、政务热线、城市治理等核心领域，并在水利、交通、文旅、医疗等创新场景持续延伸，形成兼顾效率提升与治理创新的全场景智能化矩阵。以北京市为例，据不完全统计，已有至

少 8 个政府部门完成了 DeepSeek 模型的应用部署，涉及政务服务、政务办公及城市治理等核心场景。

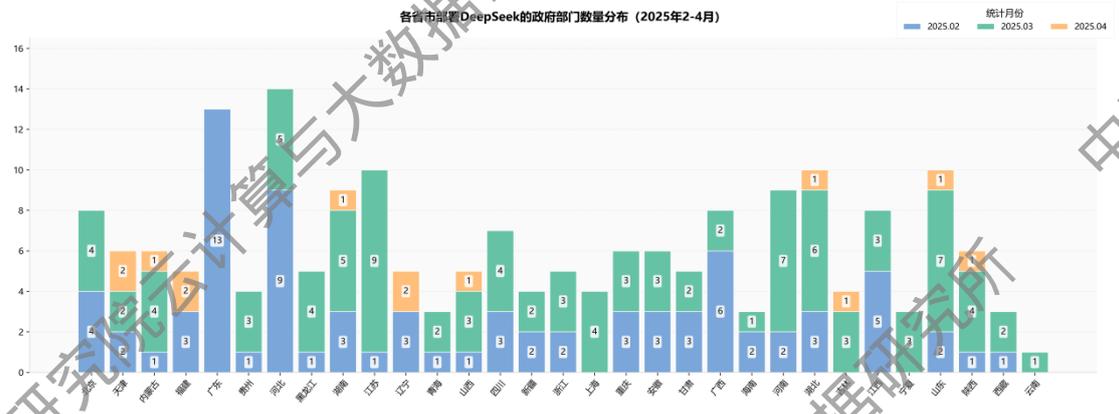
表 1 北京 2025 年 2 月~4 月 DeepSeek 应用案例

省/市	政府部门	DeepSeek 部署情况	DeepSeek 应用场景	应用类型	部署时间
北京	丰台区政务和数据局	已完成 DeepSeek 大模型在政务云的本地部署	“丰小政”数智助手聚焦政务服务“清零热线”和企业群众办事需求	政务服务	2025.02
	丰台和义街道	完成 DeepSeek 大模型本地化部署	“和小智”全面升级，政务办公智能化实现平台联动，打造治理新范式	政务办公 城市治理	2025.03
	昌平区政务和数据局	成功部署上线全参数优化版 DeepSeek-V3DeepSeek-R1 政务大模型	社区工作者可简便获取政策解读、数据分析及问题解决方案	城市治理	2025.02
	经济技术开发区 区营商环境建设局	基于 DeepSeek 大模型开发的“智能政务小助手”研发完成	“智能政务小助手”进行智能材料预审	政务服务	2025.02
	北京市市场监管局	携手 DeepSeek 大模型打造的一套集语义理解、自主学习、智能推理于一体的 AI 政务助手	“市监小 e”可提供全流程指引及注意事项，全程无障碍拟人化交互；“25+”大类业务领域智能体模型，确保咨询建议与最新政策“零时差”同步	政务服务 市场监管	2025.03
	平谷区政务服务 中心	完成“一网通办”一体化平台 DeepSeek 大模型能力的全面注入	对内综窗 AI 助手和对外自助服务一体机 AI 助手	政务服务	2025.03
	门头沟区政务服务 管理局	突破性地 DeepSeek 大模型深度融入政务服务体系	“门小政”政务平台全面升级，提升政务咨询系统效率，群众办事效率	政务服务	2025.03
	昌平区城市管 理委	智慧城市交通综合管理平台成功接入 DeepSeek 大模型	利用人工智能赋能城市管理、智慧交通治理	城市治理	2025.02

来源：公开资料整理

从部署省份覆盖广度看，全国省级行政区 100%覆盖，体现政务大模型战略布局的全面性与均衡性。DeepSeek 的部署范围已实现全国 31 个省级行政区全覆盖，从东部地区的北京、广东，到西部边疆的新疆、西藏，以及中部地区的河南、湖南，区域覆盖呈多层次、无盲点特征。东部沿海与中西部省份部署数量接近，但东部沿海地区呈现出多场景、深融合的应用生态。此外，多地区通过“AI+文旅”、“AI+基层”、“AI+安全”等特色场景切入，如黄山 AI 旅

行助手、新疆克拉玛依市“基层治理 AI 社工数字人”、昌平“回天大脑”接入 DeepSeek 应用、宜兴市“天机镜”城市安全生命体 AI 大模型，探索差异化路径，展现技术下沉的灵活性与适应性。



来源：公开资料整理

图 1 各省市部署 Deepseek 的部门数量分布图 (2025.2~4)

从省市部署密度差异看，部分省份集中效应显著，区域需求与资源禀赋双重驱动部署密度。河北、广东、湖北、山东、江苏位居前列，合计占比超 30%，反映出经济强省与人口大省对智能政务的迫切需求。其中，河北以政务服务为核心，覆盖邯郸、沧州、秦皇岛等 11 个地级市，并在水利、医疗等领域延伸；广东则以深圳、广州、福田为核心，探索“数智员工”、“智慧城管”、“城市智能体”等创新模式，覆盖政务办公、城市治理、政务服务全链条。相比之下，西部省份部署量普遍偏低，但部分区域通过“单点突破”形成示范效应，如西藏昌都市部署政务服务模型，青海司法厅接入法律咨询 AI，体现“小而精”的布局逻辑。

（三）产业渗透度不断加深，创新突破频现

政务大模型应用场景覆盖多个层面，包括洞察、治理、兴业、惠

民等。其中洞察、治理更倾向于服务数字政府自身能力建设与提升；兴业、惠民更倾向于服务数字经济开拓与数字社会发展。同时，在数字生态、数字文化等领域，政务大模型也不断与各场景结合，形成有序的大模型赋能全场景，服务于数字中国全面发展要求。

一方面，数字政府、数字经济与数字社会发展相对成熟有序，易于通过通用数字化能力如服务、治理、协同进行大模型重构，成为当前政务大模型落地的典型应用；另一方面，数字生态、数字文化等也不断通过产业数字化、数字产业化产生“双向融合、共同奔赴”的良好势头。

政务服务类	公共服务类		政务运行类		城市治理类	
民众问事	教育服务	医疗服务	内部办事	会议管理	政务服务便民热线	
企业谋策	校园助手	患者服务	自动审批	议题生成	坐席服务智能增强	
事项推理	教学定制	临床医疗	事项流转	纪要生成	人机协同决策支持	
边聊边办	教务服务	健康宣教	角色派单	任务跟踪	坐席行为智能监控与优化	
材料预审	科研助手	运营管理	归档总结	材料匹配	全流程智能化执行引擎	
智能审批	文旅服务	交通服务	公文写作	智慧党建	基层治理	城市运行
无感评价	信息咨询	公路运输	公文生成	党务问答	社区问答	视频感知
智能问数	展馆讲解	铁路运输	公文编辑	活动提醒	事件上报	智能研判
进度查询	伴游导览	水路运输	素材推荐	信息管理	矛盾调解	智能问数
	个性推荐	航空运输	审阅反馈	思想引导	总结归档	方案生成
					应急管理	水利水务
					灾害监测	水文预报
					预案生成	预警发布
					简报整理	调度预演
					指挥辅助	预案执行

来源：公开资料整理

图 2 政务大模型先锋实践场景图

从当前 DeepSeek 部署应用类型分布看，政务服务场景占据主导，垂直领域创新应用呈多点突破态势。政务服务、政务办公、政务热线三大场景合计占比 77%，构成政务智能化的基础支撑。其中，政务服

务场景以“智能问答”“政策解读”“事项推理”“边聊边办”“材料预审”“智能审批”“无感评价”“智能问数”等为核心功能，例如北京丰台区“丰小政”数智助手、贵安新区的综合性智能政务服务平台，均通过流程优化提升办事效率。此外，垂直领域创新应用占比 16%，覆盖水利（河北防汛调度“百科全书”、湖北汛情模拟推演）、交通（江苏昆山“流量预知+智能巡检”、安徽“数智执法”）、文旅（天津和平区“AI 虚拟导游”、湖南张家界“旅游市场监管 AI”）等多个领域。此外，“满血版”DeepSeek 模型多集中于复杂场景（如上海松江区基于昇腾算力的“情形引导—材料核验—跨网申办—进度追踪”的完整闭环服务体系、湖北水利厅防汛决策系统），体现技术能力与场景难度的深度匹配。

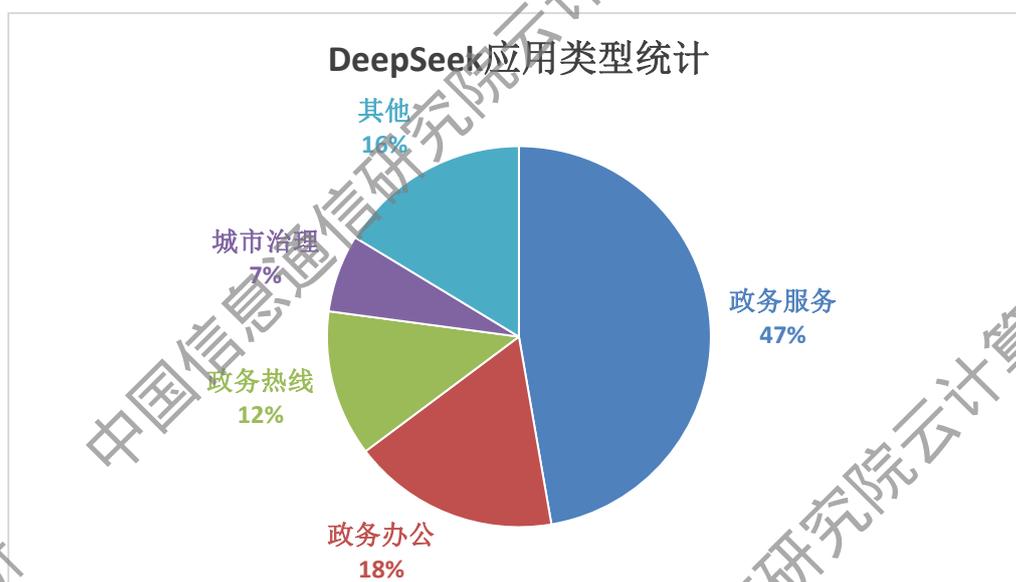


图 3 DeepSeek 政务部署应用类型统计图

二、技术流程并重，政务大模型落地建设“规致有状”

政务大模型作为数字政府建设的创新动力，推动数字治理向“好办”“智理”加速转型。大模型的场景挖掘与实践验证涉及复杂的场景需求挖掘、资源适配规划、技术可行性验证及运维保障等环节，需要充分的跨部门协作进行保障。

流程化的协作机制需要围绕“规划—论证—实施—保障”等关键环节构建，充分联动需求侧主管单位、供给侧技术服务商以及生态侧咨询单位，形成场景挖掘的先锋“合力”。

表 2 政务大模型落地实践重点工作表

阶段角色	规划	验证	实施	保障
需求侧	场景需求规划	验证指标规划	多部门业务技术协同	持续运营、能力外延
供给侧	技术方案供给	技术支持、方案供给	模型部署、调优、应用搭建、上线	持续运维
生态侧	可行性论证	选型测试	建设成熟度评估	标准化沉淀

来源：公开资料整理

（一）需求端：场景挖掘上，强调多部门协同共致

1. 规划阶段需求挖掘应切中“痛点堵点”

场景需求的发掘应从当前流程痛点或行业发展重点入手，由主管委办局牵头提出。一是要建立科学化、结构化的需求采集机制，确保场景需求与业务痛点精准匹配，为后续资源统筹和技术适配奠定基础。二是探索清单提报机制，系统性梳理各部门需求，避免技术应用与业务需求脱节。从地方实践来看，各地普遍采用由相关职能部门牵头提出需求，并通过行业数字化行动计划或实施意见等形式呈现。例如，

《上海交通数字化转型实施意见》由上海市交通委员会印发，其中明确提出交通大模型建设需求。此外，2025年3月，青岛市发布首批30个“AI+政务”赋能应用场景清单，覆盖政务服务、城市管理、政务运行、医疗健康等领域。这一机制不仅加速了供需对接，还通过开放63项需求清单（含31项市直部门和32项区级部门需求），推动技术与企业、市民需求的深度融合，加强结构化需求提报在激发市场活力与创新转化中的关键作用。

2. 验证阶段方案论证应兼顾“内修外道”

需求验证应兼顾本地落地可行性与转化潜力，充分借鉴其他地市、其他需求场景的落地经验。

一是围绕核心业务需求构建验证框架，重点涵盖业务场景分类、应用成熟度分析、数据资源规模、系统性能阈值、架构耦合度评估及安全合规要求六大核心维度，通过结构化指标体系实现需求侧信息的系统化采集与量化分析。

专栏 1：政务大模型硬件资源需求评估

政务大模型模型参数与运行参数的合理配置是资源审核的核心前提需基于场景复杂性、敏感性分级制定技术方案。在政务大模型私有化部署中，模型参数（满血版或蒸馏版）与运行参数（上下文长度、批次大小）的协同设计直接决定算力需求规模。对于涉及复杂业务流和应用系统以及复杂业务知识及逻辑的场景，可采用满血版模型并进行微调构建行业应用场景或通过强化学习构建专属大模型。而在语言类的政务咨询、情感分析、文本翻译和图像类检测、识别等一般性场景中，可采用 7B-13B 的蒸馏版模型，通过参数精简降低算力消耗，同时满足政务服务的响应时效要求。运行参

数的设定需结合政务业务的潮汐式特征，动态调整批次大小与序列长度。例如，在民生政策咨询高峰时段，可通过增加批次大小提升并发处理能力，而在非高峰时段则优化序列长度以降低能耗。同时，对于有长文本如政策文件解析等场景需求的场景，可配置 32K 以上的上下文长度。

政务大模型资源审核需围绕显存容量、算力性能展开，通过精准匹配模型需求与硬件参数，确保政务场景下推理效率与资源投入的全局最优解。显存容量是政务大模型部署的硬性约束条件，由基础模型占用、运行时缓存及系统开销三部分共同决定。

显存容量是政务大模型在硬件选型的首要约束条件，而计算能力与带宽参数决定推理性能上限，需结合并发需求与成本预算综合决策。计算能力作为算力核心指标，由芯片架构与软件栈协同优化的实际浮点算力（如 INT8、FP16 等）决定，直接影响单卡推理吞吐量（Token/s）的理论上限；显存带宽则制约计算单元与存储介质间的参数加载效率，当批量处理规模突破带宽承载阈值时，系统性能将受限于数据传输速率。在多卡分布式部署场景中，互联带宽进一步成为关键瓶颈，决定计算节点间梯度同步与数据交互的时效性，影响集群算力扩展的线性比例。硬件部署需综合平衡显存供给、算力密度与带宽资源的匹配关系，结合政务场景对并发响应能力、安全冗余等级及成本投入的约束条件，通过模块化资源配置实现推理效率、服务规模与基础设施经济性的动态优化。

以部署 DeepSeek-R1 671B 模型为例，计算所需使用昇腾 910B（显存容量：64GB）显卡的张数。

- 模型权重大小是从满足模型部署的最低要求，首先需要考虑显存容量是否足够。模型权重大小占用显存为模型参数量与计算精度的乘积，如 671B 参数模型采用 1 字节（INT8 精度，昇腾卡不支持 FP8 需转到 INT8 类型）时占用 671GB；

- 激活值缓存，模型运行时产生的中间计算结果，与模型参数和精度相关。计算方式为模型参数量、精度及动态系数的乘积，如 $671B \times 1 \text{ 字节} \times 0.25 = 167.75GB$ ；
- 输出张量缓存，模型生成结果所需的临时存储空间，则与批次（模型一次处理的请求数量）、序列长度及词表规模（模型能够识别和处理的不同单词或标记的数量）正相关，如 $16 \text{ 批次} \times 8192 \text{ 序列} \times 128256 \text{ 词表} \div 1024^3 (1B \rightarrow 1GB) \approx 15.66GB$ ；
- 同时预留 1GB 固定开销用于系统初始化，包括软件栈缓存、算子编译缓存等。

据此，若部署 671B 大小模型，则单实例显存总需求约 855.41GB，至少需要 14 张华为昇腾 910B 实现模型部署最低要求。然而，政务场景的特殊性需显存配置预留 20%-30% 冗余容量，以应对高并发公共服务峰值压力，如千人级在线咨询，及多级安全隔离需求，如开发/生产环境物理分割，确保系统在高负载与复杂安全策略下的稳定运行。

二是建立由技术专家、行业顾问及业务部门构成的联席评审机制，依据战略性、可行性、协同性三重标准对需求清单开展分级评估，充分吸取多维的评估意见。

三是充分推动资源集约化，通过业务流程映射和技术架构比对，精准识别系统集成接口、数据共享节点及基础设施复用机会，形成需求优化方案与实施路径规划，确保资源投入的精准性和建设效能的集约化。

四是广泛借鉴成功经验，各地在探索政务大模型场景应用过程中已积累了不少成功案例，如宜兴市完成 DeepSeek 部署并落地热线坐席助手等多个政务应用、武汉云上线“满血版” DeepSeek-R1 模型，江城智能政务应用再升级等。通过典型案例分析，提炼出适用于本地的共性场景和可复制模式。例如，某地在政务服务中采用大语言模型

+RAG 技术实现“智能问答”，其他地区可据此判断是否具备相似业务场景及数据基础。

专栏 2：政务大模型建设方式与资源选择

政务大模型部署需兼顾技术效率与业务需求，通过构建多维协同、灵活适配的部署模式，推动数字政府服务体系升级。建议结合区域实际情况进行选择：

一是部省联动，强化中央行业主管部门与省级政府的协同，聚焦行业垂直领域（如公安、医疗、教育）共建大模型，由部层面统筹行业标准与数据规范，省级负责区域特色数据融合与模型训练，形成“行业纵深+区域特色”的双轮驱动模式，破解跨层级数据共享难题。

二是地方统建，地方主管部门主导横向打通各部门数据壁垒，集约建设统一大模型基座，采用“厚基座+薄应用”架构，为政务场景提供标准化模型服务，实现算力、算法、数据的全省共享与高效利用，避免重复建设。

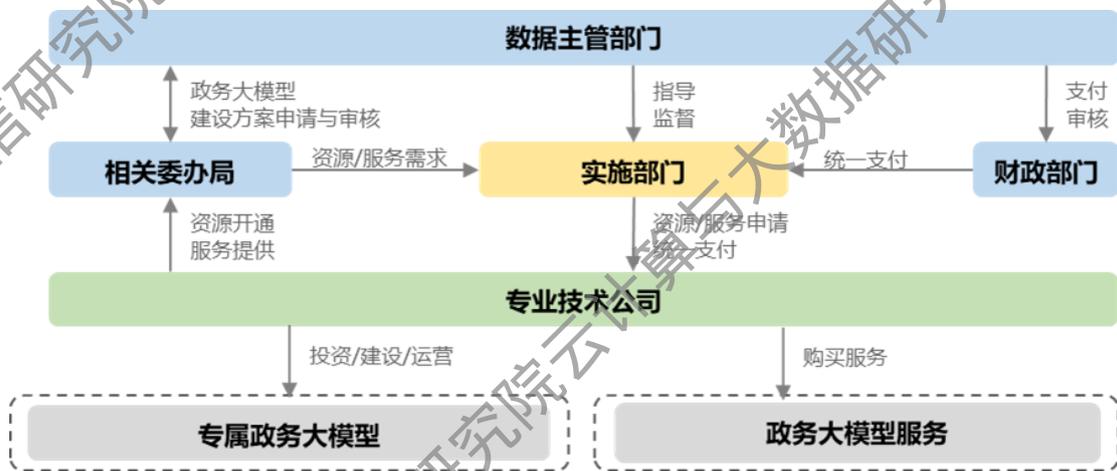
三是分级建设统一管理，建立省、市、县三级联动机制，省级统筹模型技术框架与资源调度，市县级聚焦本地化需求开展场景适配与增量训练，形成“全局统一规范、局部灵活创新”的分级共建模式，平衡规范性与灵活性。

四是互联网直接调用，面向高频民生服务场景（如社保查询、证照办理），通过安全接入互联网通用大模型能力，构建轻量化政务智能客服，降低开发成本，提升服务响应效率，但需建立严格的内容审核与数据安全防护机制，保障服务合规性。

3.实施阶段部门协同应推动“有序协作”

在相关数据主管部门授权和统筹指导下，由相关委办局和实施部门协同细化政务大模型落地实施过程，共同指导专业技术公司完成相关服务的部署、开通、调优。

- 数据主管部门职责：数据主管部门负责政务大模型项目建设方案审核审批、汇总项目库，运营指导监督，支付资金审核。
- 相关委办局职责：各级政府业务部门提出政务大模型建设业务需求和资源需求。
- 相关实施部门职责：相关实施部门负责数字化整体运营，统筹相关资源、服务申请，对内提供统一服务。
- 财政部门职责：财政部门负责编制年度预算，支付资金。



来源：公开材料整理

图 4. 政务大模型项目实施协作关系图（需求侧）

4.保障阶段运营规划应注重“生态培育”

政务大模型持续运营需聚焦数据、模型与场景的动态协同进化，构建自优化的生态体系。在大模型场景建设完成后，运营管理的核心目标是确保技术的持续迭代与效能优化。通过引入统筹的运营团队服务，建立敏捷的技术更新机制，依托统一的算力调度平台和知识中枢，实现算力资源的灵活调配与知识库的动态更新。充分发挥市场主体的创新能力，同时通过弹性投入机制降低财政压力，推动技术升级的可持续性。以数据闭环驱动模型能力提升，通过常态化的数据采集、标

注与质量评估机制，持续注入高质量政务领域新知识、新规则，确保模型知识图谱的时效性与准确性；建立模型动态调优机制，基于实时业务反馈与用户交互数据，通过微调、强化学习等技术手段优化模型参数与响应策略，提升政务问答、智能问数等场景的适配性。同时，深化场景拓展与用户需求洞察，通过跨部门协作挖掘政务服务痛点，推动模型从单一问答向政策解读、风险预警、流程自动化等多元化应用延伸，构建“数据迭代-模型进化-场景深化”的正向循环，实现大模型对政务效能提升的持续赋能。

（二）供给端：技术服务上，强调全过程闭环落地

面向供给侧，在提供政务大模型的落地服务过程中，应聚焦全过程闭环落地，因地制宜提供技术支持。



来源：公开材料整理

图 5 政务大模型落地服务关键步骤图（供给侧）

1. 规划阶段方案设计应聚焦“软硬协同”

在方案设计过程中，服务商应充分结合软硬件实践基础，立足场景需求，推动业务与技术的高效协同。在顶层规划阶段，需立足政府数字化转型目标，明确大模型赋能的核心方向，例如通过政策智能匹

配优化产业扶持效率、依托民生诉求挖掘提升公共服务精准度等；战略层面需考虑到技术、业务、管理多方面，统筹政策合规性（如数据安全、AI 伦理规范）、资源整合路径（如政务云底座复用、跨部门数据资产目录梳理）及组织协同机制（如成立专项工作组）；解决方案设计需以“轻量化切入、模块化扩展”为原则，优先搭建高兼容性技术中台（如支持多源异构数据接入的预处理引擎、可插拔的领域微调模块），并同步制定数据治理标准（如政务知识图谱构建规范、敏感信息脱敏流程）；可行性论证需综合考虑到技术、经济、风险等维度，量化大模型在算力消耗、响应时效、人机协同成本等方面的投入产出比，并建立动态风险防控机制（如 AI 决策可解释性验证、偏见监测与修正）；业务场景筛选则需基于需求紧迫性、实施可行性、价值显性度等方面，锚定如“一网通办”智能导办、基层治理事件智能分拨等标杆场景，形成分步落地的路线图。

2. 验证阶段方案迭代应做到“小步快跑”

验证场景 POC 需以“小步快跑、快速迭代”为原则，构建技术能力与业务价值的双向反馈闭环。在 POC 阶段，建议选取 3-5 个高颗粒度场景（如企业证照变更智能预审、12345 热线诉求自动归类），搭建包含数据沙箱、模型训练舱、效果评估看板的一体化验证平台。

技术验证：覆盖多维度能力测试，包括模型对政务专业术语的理解准确率（如政策文件中“负面清单”等概念的语义解析）、长文本多轮对话的上下文保持能力（如企业补贴申报咨询场景的连贯性响应），因果推理和复杂逻辑推理能力（如建设项目审批条件交叉校验）

以及分业务场景的差异化输出能力（如政务服务和政务办公场景的对话人角色不同）。

技术选型：通过 AB 测试对比开源模型微调、行业大模型调用、混合增强架构等不同方案，结合政务专属知识注入效率（如地方性法规快速适配能力）、系统扩展成本（如 GPU 集群扩容边际成本）等维度进行综合优选。最终形成包含场景适配度评分、风险清单、规模化推广阈值的 POC 验证报告，为全面落地提供决策依据。

场景效果评估：建立“业务指标+技术指标”双评价体系，例如在舆情预警场景中，既需验证预警准确率、响应时效等技术参数，也要评估其对基层人力成本的节约效果。

3.实施阶段项目建设应助推“规范高效”

在项目实施过程中，服务商应充分立足于建设单位业务情况，按需推动相关关键技术底座建设。在相关数据治理平台、训练开发评价以及原生应用开发引擎等技术底座构建过程中，应参考相关标准规范，推动技术底座与上层业务的高效协同。

3.1 数据治理平台建设

政务大模型数据治理平台应构建全模态、全链路的一站式数据工程能力，为大模型训练提供高价值、高安全性的政务数据资产。平台建议覆盖从数据获取到服务供给的全生命周期，通过标准化、自动化、智能化的技术架构，实现政务数据资源的深度治理与高效利用，从而支撑政策分析、民生服务、城市治理等场景的智能化升级。

多模态数据整合能力是平台的核心基础，应实现多源异构政务数

据的统一接入与结构化解析。平台宜支持文本(政策文件、信访记录)、图像(证照扫描、监控画面)、音频(热线录音)、视频(会议记录)等多类型数据的批量与实时接入,兼容 OBS、本地存储、数据库等异构数据源。针对非结构化数据,内置 OCR、语音识别、视频关键帧提取等预处理工具,将原始数据转化为标准化结构信息,例如从会议视频中自动提取发言摘要并关联议题标签,构建可分析的语义化数据资产。

智能化数据预处理管线应深度融合政务领域知识,并结合应用目的实现数据的精准提纯与语义增强。平台可考虑内置政务专用算子库,包括政策条款提取、地址归一化、实体提取、敏感信息脱敏等场景化处理功能,并通过低代码界面支持算子灵活编排。例如,针对“企业开办”业务场景,可配置“证照识别→经营范围分类→合规性核验”自动化清洗链路。同时集成大模型能力,实现非结构化文本的意图识别(如区分咨询与投诉)、图像数据的实体关联(如将监控画面中的异常事件映射至网格化管理体系),显著提升数据语义价值。

人机协同标注体系应兼顾效率与安全,支撑政务场景下的高质量知识沉淀。推荐平台支持多模态数据的智能标注能力,包括文本 QA、图片 QA、问答排序、图片物体识别等标注任务,可通过 AI 辅助标注、团队协作标注等模式提高标注效率与准确性。政务数据涉及敏感信息,平台宜建立严格的标注安全机制,例如对音频、视频数据进行片段切分与乱序处理,确保标注过程中敏感信息不被泄露。同时,平台鼓励支持数据集与标注任务的解耦设计,实现标注任务的灵活配置

与数据集的独立管理，便于根据不同大模型训练需求快速构建适配的数据集。

数据质量闭环管理机制应实现治理流程的自优化与自迭代。平台需考虑质量检测、根因分析、策略调整等方面进行自动化链路设计，基于规则引擎与 AI 模型实时监测数据完整性（如字段缺失报警）逻辑一致性（如户籍地与社保缴纳地冲突检测）等核心指标。针对高频问题数据（如格式混乱的基层报表），自动触发强化清洗策略或通知数据源部门整改。通过可视化质量看板呈现数据健康度指数，辅助决策者优化资源配置。

多层次安全防护体系应贯穿数据全生命周期，构建可信政务数据流通基座。平台可通过构建多层次的安全防护体系，覆盖数据获取、存储、加工、标注和流通等全生命周期环节。针对不同分类分级的政务数据，平台鼓励实施差异化的安全策略，例如对核心政务数据采用更高的加密级别、更严格的访问控制和更频繁的审计机制。同时，建议平台支持数据血缘分析功能，清晰记录数据的来源、流转过程和使用情况，确保数据的可追溯性与合规性，为政务数据的共享应用提供安全可信的基础。

场景化数据服务能力应精准对接大模型训练需求，实现治理成果的价值转化。该服务能力宜支持将处理后的高质量数据集供给预训练、微调、RLHF 等大模型训练场景。平台需根据不同训练场景的需求，提供多样化数据集格式与内容，例如为微调场景提供单轮问答库、带角色单轮问答库、多轮问答库等，并确保数据集的版本管理与动态更

新能力。特别是针对 RLHF 训练场景，推荐平台支持构建包含人类反馈的多模态数据集，例如政策解读对话对、公共服务多轮交互记录等，并通过标准化的反馈收集流程与多轮验证机制，确保反馈数据的一致性与可靠性。

通过工程化、智能化的数据治理体系，该平台将成为政务大模型的核心基础设施。以“城市应急响应”场景为例，平台可实时整合气象、交通、社情等多模态数据，经清洗标注后生成事件演化图谱，驱动大模型输出资源调度方案与风险预警建议，助力政府构建“感知-决策-处置”的智能治理闭环，全面提升公共服务效率与城市韧性。

3.2 大模型训练开发平台建设

政务大模型训练开发平台应构建模型全生命周期管理体系，成为数字政府智能化转型的核心技术基座。平台宜覆盖模型开发、训练、评测、压缩及推理全流程，通过标准化工具链与智能化服务能力，满足政务场景对安全性、合规性及高效协同的严苛需求，助力政府实现 AI 技术的深度应用。

云原生算力调度体系应实现异构资源的高效整合与弹性供给。建设异构算力调度平台，整合政务算力资源，形成统一推理算力资源池。采用服务租用模式，提供一体化、弹性可扩展的算力服务。基于容器化与动态资源分配技术，平台鼓励支持 CPU、GPU、NPU 等多元算力的统一纳管，通过智能调度算法实现训练、推理、边缘算力的按需分配。针对突发性政务需求，推荐建立算力资源池化机制，宜支持分钟级弹性扩容与成本最优调度策略，确保关键任务优先级与资源利用

率双达标,支撑大模型训练与推理需求得到及时响应。

全栈模型训练工具支撑灵活化训练需求。平台可通过集成高性能模型训练工具,支持预训练、全量监督微调(SFT)和 LoRA 等轻量化微调技术,为不同规模和复杂度的政务模型提供灵活的训练方案。内置的训练框架需兼容主流深度学习框架,并提供分布式训练优化能力,确保大规模模型训练的效率与稳定性。同时,平台鼓励具备自动化模型评估功能,通过多维度指标(如准确率、响应时间、资源消耗)对训练过程进行实时监控和分析,为模型优化提供数据支撑。

多维度模型评测体系保障性能透明化验证。建立大模型应用评测管理平台,提供内容质量、合规性、安全性及性能评估服务。加强生成内容审核与数据调用审计,防范模型“幻觉”、虚假信息等风险。通过评测后正式上线服务。模型评测环节需构建覆盖分类准确性、回归误差及生成质量等多维度的评估体系,通过可视化界面直观展示评测结果。平台可考虑支持用户选择特定数据集进行评测,并提供混淆矩阵、ROC 曲线、Loss 曲线等图表形式,帮助用户全面分析模型在不同场景下的表现。同时,平台可集成标准化评测流程,支持评测结果导出为报告,便于跨团队协作与持续改进。通过量化指标与定性分析的结合,确保模型性能验证的客观性与可追溯性。

模型压缩技术赋能边缘场景高效部署。针对大语言模型的推理性能优化,平台通过引入量化压缩技术(如 INT8/FP16 低精度表示法),在降低模型参数存储需求的同时保持较高预测精度。此外,推荐平台支持剪枝技术与知识蒸馏算法,进一步缩小模型体积,提升边缘设备

上的部署效率。这些优化措施不仅能够减少硬件资源占用，还能显著缩短推理延迟，从而增强模型在资源受限场景下的实用性与用户体验。

3.3 模型服务平台

模型服务平台实现模型资源的统一调度、服务生命周期的智能监管及内容安全防护。该平台通过模型服务 API 整合多样模型资源并提供安全可信的调用能力，依托模型服务监管实现从模型部署到应用分析的全流程监控与治理，并借助模型服务安全构建覆盖输入、处理与输出的动态防护体系。三者协同保障模型服务平台的高效、合规与可持续运行，为政务智能化应用提供坚实支撑。

模型服务 API 旨在提供统一、安全且可运维的多样化大模型调用能力。模型服务 API 可通过整合多家服务商的不同参数规模模型，形成灵活的模型资源池，并支持统一的模型 API 网关实现流量控制、权限管理及参数配置。模型服务 API 宜建立用户权限管理功能，确保用户仅能访问其权限范围内的模型，并集成数据加密、访问控制等安全机制。模型服务 API 建议支持完整记录 API 请求与响应日志，并具备实时监控功能，可跟踪模型性能、错误及资源使用情况，便于问题排查与性能分析。模型服务 API 推荐支持模型服务的上线、下线操作以及简化的部署流程，并能根据流量需求动态扩缩容服务。此外，模型服务 API 可为后续政务 AI 原生应用引擎智分配所需模型服务，并对其运行使用情况进行管理、监控及统计分析等运营操作。

模型服务监管实现模型从部署到应用的全流程管控与效能优化。模型服务监管宜支持基础模型监管工具服务，通过监控可用性以及核

心性能指标，保障模型运行稳定性；同时依托应用监管工具服务，深入用户行为监控、访问记录分析、知识库调用追踪及活跃度评估，结合 API 调用统计、输出日志分析与操作审计等，实现模型使用过程的透明化与风险预测；在此基础上，通过模型市场的入驻、审核、上架、下架等全生命周期管理机制，以及大模型服务中心和通用 API 服务中心的统一展示与对接支持，形成从模型纳管到服务交付的标准化、可视化运营闭环，最终支撑政务场景下模型服务的规范化、高效化与持续迭代。

模型服务安全构建贯穿输入、处理与输出全流程的内容审核与动态响应机制。模型服务安全可通过多语种风险识别服务对涉政、违法等敏感内容进行输入审核，结合 Prompt 审核服务实现对提示词注入、不良价值观等异常行为的多维度检测。依托红线知识库服务为涉政等特殊议题提供标准化回复模板，确保内容合规性，同时宜建立回复干预服务，针对突发安全事件提供语义、文本、关键词等多层级应急处置能力。在输出端推荐配套输出安全检测服务，通过兜底回复、不上屏等策略保障生成内容合规性，并集成输入输出安全内容更新服务，支持策略模板管理、安全词表配置及白名单输入控制，实现动态策略适配。此外，模型服务安全可考虑实现输入输出日志的全链路记录、分析与存储，结合内容溯源能力确保风险事件可回溯、可审计，最终形成从风险识别到动态防护的闭环安全服务矩阵。

3.4 AI 原生应用引擎

政务 AI 原生引擎赋能政府大模型全场景落地，打造 AI 应用创新

基座。平台通过模型选型、知识治理、智能体开发与应用服务的深度整合，实现大模型技术与政务场景的深度融合，赋能政务咨询、民生服务、政务办公等核心领域智能化升级。

模型层需构建多样化能力矩阵，支撑复杂政务场景的智能决策需求。制定大模型选型标准，遴选并部署主流开源与闭源模型，形成涵盖通用大模型、多领域垂直大模型和共性基础工具的区域模型矩阵。在模型层，政务 AI 原生应用引擎强调集成多样化的模型选择，这包括但不限于开源与闭源模型，以及基础模型与行业大模型，旨在构建一个全面的模型仓库，以增强政务智能体对复杂世界知识的理解和处理能力。多样化模型的提供为政务智能体提供强大的思考能力、推理能力、安全能力和检索能力，这些模型能够处理复杂的业务问题，提高决策的精准度。

政务知识库需构建领域知识中枢，为 AI 决策提供精准可靠的知识支撑。搭建大模型知识中枢平台，实现政务领域知识的汇聚管理。提供知识治理共性工具，支持多模态数据的收集、处理与更新，为构建推理知识库、训练数据集与模型提示词提供支撑。平台宜整合政策法规、历史档案、业务案例、研究报告和最新政务动态等多元数据源，建立覆盖知识采集、加工、存储、应用的全链路管理体系。通过智能文档解析技术实现非结构化数据的语义化处理，结合知识图谱技术构建实体关系网络，形成可检索、可推理的知识体系。在知识应用层面，集成 RAG 框架实现与大模型的深度协同，通过动态知识检索增强模型输出的准确性，结合 Prompt 工程优化任务指令的领域适配性，形

成知识驱动智能决策的完整闭环。

政务智能体技术通过融合实时感知、智能规划、动态记忆与精准执行四大核心能力，为政府业务场景提供智能化支撑。其实时感知模块是通过政务平台、物联网设备、新闻媒体等多源数据实时采集，动态感知民生诉求、城市运行状态和最新形势（如 12345 热线诉求热点、交通拥堵指数、国内外最新形势）；规划模块采用任务分解与多算法协同机制，将复杂政务任务拆解为可操作的子任务序列，该架构突破传统线性流程限制，支持非结构化任务的动态编排能力；记忆系统构建分层存储模型，短期记忆模块实时捕捉业务流程中的交互数据，长期记忆库整合政策法规、历史案例等结构化知识，结合语义检索与向量化技术实现跨模态信息精准匹配。执行层创新性地引入模型上下文协议（MCP），建立智能体与政务工具的标准交互范式。通过协议层对数据库接口、云服务平台等异构系统进行统一抽象，实现工具能力的动态编排与线性扩展。

政务智能体开发平台构建端到端的敏捷开发体系，支持可视化定义政务角色画像与 workflow 逻辑。开发者可通过拖拽式界面配置政策解读、民生服务等场景的专属处理流程，集成智能工单分类、知识检索等定制化工具链。该模式既保持核心政务流程的标准化，又赋予区县层级灵活调整服务策略的空间，实现政务服务统一性与区域适应性的平衡。此外，在政务智能体评估与优化阶段，系统通过任务准确性评估、工程性能评估以及安全评估，全方位检测智能体的表现。例如，优化 AI Agent 构建流程对于提升开发效率起到了显著作用。最后在

确保数据安全的前提下，系统支持 API 网关与 Web 服务的多渠道发布，满足窗口终端、移动政务等多场景接入需求，形成闭环的服务质量保障体系。

4.保障阶段持续运维应确保“合规可靠”

服务商的政务大模型运维体系以确定性为核心，构建全链路风险防控与快速响应机制。通过全生命周期主动运维策略，深度融合故障模式库与深度巡检能力，结合混沌工程模拟极端场景，系统性识别潜在风险并提前消减漏洞，形成覆盖硬件层、模型层与应用层的防御闭环。基于智能监控与告警联动机制，实现异常行为的毫秒级感知与精准定位，同步触发自动化隔离与流量调度策略，确保故障对业务连续性的影响最小化。运维合规性贯穿资源调度、权限管控与操作审计全流程，通过标准化流程与动态策略优化，推动运维模式从被动处置向主动防御转型，为政务服务的稳定性与安全性提供底层支撑。

服务商应通过多层次防护体系与动态治理机制构建安全防护体系，确保政务智能化应用的合规性与可靠性。其核心技术涵盖数据全生命周期安全管控，包括数据来源合规审查、敏感信息加密脱敏、标注流程规范及细粒度访问控制，保障全流程数据主权控制确保数据信息流转受控、权责可溯；模型安全层面通过训练过程审计、参数加密存储、分类分级管理及备案登记制度，防范模型泄露与篡改风险；内容安全体系采用双轨审核机制，可结合语义分析与人工复核对输入输出内容进行敏感词过滤与价值导向校验；运行监测系统依托监管沙盒实现风险预判，通过实时流量追踪与异常行为分析构建主动防御能力，

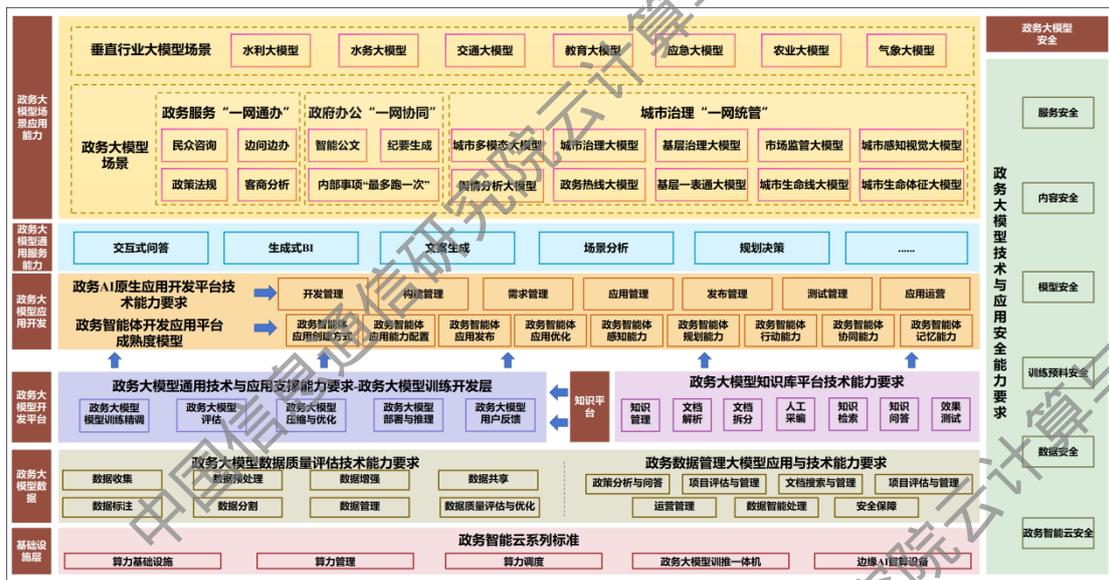
并建立应急响应链路快速处置数据泄露或 API 攻击事件；技术底座整合纵深防御架构与可信供应链管理，强化身份认证、接口防护及开源组件漏洞修复能力，形成从基础设施到上层应用的全链路安全屏障。通过安全与智能的深度耦合，既拓展了政务服务智能化转型的可行边界，也为构建可信的数字政府提供可验证的技术支撑路径。

（三）生态端：行业生态上，强调标准化以评促优

政务大模型标准化建设是技术可靠性与服务价值落地的全周期管理基座，为成效评估到迭代优化闭环提供可靠支撑。第三方机构应围绕政务大模型的建设过程、实施过程、应用成效、运营运维、安全保障等方面，推动政务大模型发展“有据可依”。**建设过程标准化**通过规范技术选型、底层算力基础设施搭建、数据源合规性及算法设计逻辑，确保模型开发与政策目标、业务需求深度耦合，规避技术路线偏离导致的资源错配；**实施过程规范化**以统一协作流程与资源调度规则降低跨部门协同成本，提升模型落地效率与规模化复制能力；**应用成效可量化**依托客观评估指标验证模型对政务服务效能的提升效果，避免技术投入与公共价值脱钩；**运营运维持续化**通过动态监控、迭代优化及故障响应机制保障系统长期稳定性，适应业务场景的动态变化；**安全保障体系化**则从数据隐私保护、算法抗攻击性及系统韧性层面构建全域防御体系，防范技术滥用与安全漏洞引发的公共治理风险。上述维度共同构成覆盖全生命周期的闭环管理框架，既确保技术落地的可控性，又强化公共服务输出的可持续性，是政务大模型实现可信、

可用、可扩展的核心基础。

前期，中国信通院云大所围绕政务大模型方向，推动标准制定工作。当前已构建政务大模型标准体系 2.0 版，其架构覆盖政务智能云、政务大模型一体机等底层基础设施，贯穿数据治理、知识平台构建、模型开发与智能体、AI 原生应用开发等核心技术环节，并延伸至服务能力输出、场景化应用适配以及全生命周期安全保障，形成分层递进、多模块协同的技术规范框架。同时，相关评估体系的落地和推广将进一步巩固政务大模型建设案例的引领作用，提高其可信度和可靠性，增强用户的信任感。未来，随着政务大模型标准与评估体系日臻完善，将会进一步推动政务行业大模型健康有序发展。



来源：中国信息通信研究院

图 6 政务大模型标准体系 2.0 全景图

三、四维要素共塑，政务大模型未来发展“趋势向善”

（一）构建高质量政务数据基础，增强政务大模型服务精准性

高质量数据集作为政务大模型进化的核心基础，需通过动态更新与优化、严格合规性保障以及数据的精准与一致性，推动政务智能系统的持续创新与升级。随着政务数字化转型的加速，数据来源日益丰富多元，包括政务服务平台、物联网设备、移动政务应用以及跨部门数据共享等渠道，这些数据为政务大模型提供了海量且多维度的学习资源。例如，政务服务大厅的业务办理记录与在线政务平台的用户操作数据相结合，能够为模型提供全面的用户行为画像，为精准服务推荐提供支持。为确保数据集的时效性，通过建立动态更新机制，及时纳入最新政策法规、社会经济动态等关键信息，使政务大模型能够快速适应政务环境的变化。同时，合规性是高质量数据集的基石，需严格遵循国家数据安全与隐私保护法律法规，对数据进行脱敏处理与权限管理。此外，数据的精准与一致性也同样关键，通过数据清洗、标注与校验流程，去除噪声数据，统一数据标准，确保模型输入数据的质量。通过构建动态、合规、精准的数据集体系，促进政务大模型能力提升，提升决策支持的准确性与服务优化的效能，为政务智能化转型提供坚实的数据支撑，推动政务领域应用服务迈向更深层次的智能化与人性化。

（二）智能体驱动治理体系重构，强化政务大模型决策

灵活性

智能体通过“低成本知识更新-高效工具集成-自任务闭环”等多重优势，突破传统大模型能力边界，重塑政务领域应用服务的敏捷性与可执行性。在动态知识融合能力方面，智能体通过集成向量存储检索增强生成、大模型、任务规划、工具接入、记忆功能等技术能力，可实时整合结构化与非结构化知识资产，构建可演化的专题知识库，并通过反思机制实现知识的持续优化，解决了大模型依赖预训练静态知识导致的迭代成本高、实时性不足的缺陷。在敏捷工具协同能力方面，基于 MCP（Model Context Protocol，模型上下文协议）的标准化接口，智能体可动态编排异构政务系统的功能模块，实现跨平台工具的原子化调用与并行调度，提升系统响应灵活性与功能复用效率。在全流程任务自治能力方面，智能体可自主拆解复杂目标为可执行子任务，并通过状态追踪与动态纠偏机制实现端到端的执行闭环。在此模式下，人类角色聚焦于战略目标设定与关键决策监督，而智能体完成从任务解析、资源调度到结果反馈的全链路操作，从根本上解决大模型“认知与执行割裂”的局限，推动政务领域应用服务从信息辅助向自主治理的质效跃升。

未来政务智能体的技术架构将向“协议化工具交互+分布式协作网络”方向演进，通过 MCP 协议、A2A（Agent-to-Agent）等架构实现工具生态解耦与认知动态化，重构政务智能化系统范式，支撑技术融合、业务融合、数据融合，提升跨层级、跨地域、跨系统、跨部门、

跨业务的协同管理和服务水平。

（三）牢筑全链路内生安全体系，提升政务大模型可靠性

政务大模型安全体系将迈向“模型基座-知识库-训练开发-智能体-应用服务”全链路内生安全体系，通过技术加固与动态防御机制，实现风险防控与价值对齐的双重目标。技术层面目前仍需聚焦三大核心风险：基座模型的“幻觉”输出、知识库的数据泄露隐患及智能体的越权访问漏洞。针对基座模型的安全加固，可融合安全推理链技术，通过输入意图识别、输出语义合规性校验及逻辑一致性验证的等多重过滤机制，阻断敏感内容生成并提升决策可靠性；知识库防护则可结合联邦学习与隐私计算等技术，在保证跨域数据协同的同时，防止训练数据逆向还原与隐私泄露；智能体安全则依托动态权限控制与行为审计机制，通过零信任架构实现细粒度访问管理，确保任务执行的合规性。技术架构的演进方向将强调“安全内生”特性，将防御能力嵌入模型开发、部署与运行的各环节，而非依赖外部补丁式修复，从而形成自适应、自进化的安全防护能力。

政务大模型安全治理机制将从“单点合规”向“制度-技术-伦理”三位一体生态演进，通过标准引领与对抗性防御体系，构建可持续的信任闭环。政务大模型的安全治理突破技术单点优化的局限，转向多方协同的生态化治理模式。一方面，随着《生成式人工智能服务管理暂行办法》等法规的完善，政务大模型领域将形成垂直化、细粒度的

行业标准，例如针对模型安全、数据安全、服务安全等领域制定专项安全合规指南，明确数据采集边界、内容生成伦理等关键要求。另一方面，安全生态的构建需要政府、技术厂商、第三方机构等多主体深度协作，政府主导制定安全战略与监管框架，技术厂商研发适配政务场景的专用安全工具，第三方机构则提供独立测评与持续监测服务。此外，通过建立覆盖“数据-模型-应用”的全链路安全评测体系，结合自动化内容审核与人工复核机制，实现从训练语料筛选到生成内容风控的全流程治理，最终形成“政策牵引、技术赋能、生态共治”的可持续安全格局，为政务大模型的合规化、可信化发展提供系统性保障。

（四）创新跨部门组织协同方式，强化政务大模型服务敏捷性

政务大模型落地涉及跨部门协同，数据治理、技术开发、场景梳理等复杂要求，传统组织模式面临挑战。一是部门壁垒阻碍数据共享。垂直管理体系导致数据孤岛长期存在。二是部门能力与需求错位。业务部门和技术部门在知识结构和认知上存在差异，导致对大模型落地的步骤、过程和难点存在偏差。三是响应时效或难以支持技术节奏。按照传统审批流程耗时较多，与大模型“快速迭代”需求可能脱节。

政务大模型落地不应局限在技术创新和业务创新的范畴，组织方式上也应积极探索。一是成立政务大模型联合专班，实现跨域协同。由数据局牵头，整合业务部门、技术企业、高校专家组建实体化专班，

建立“需求对接-开发测试-上线推广”全流程通道。此模式打破部门间原有的协作壁垒，建立“业务需求-技术实现”直通机制，促进信息共享，提升决策效率，缩短开发和迭代周期。二是建立政企学研联合实验室，共建生态。