ICS 35.240.99
CCS L67

团　　体　　标　　准

T/ISC XXX—XXXX

政务大模型通用技术与应用支撑能力要求

General technical and application support capability requirements for government large models

在提交反馈意见时，请将您知道的相关专利与支持性文件一并附上。

征求意见稿

2025-03-28

XXXX－XX－XX 发布　　XXXX－XX－XX 实施

中 国 互 联 网 协 会　　发布

目　　次

前　言

本文件按照GB/T 1.1—2020《标准化工作导则　第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中国互联网协会提出并归口。

本文件起草单位：中国信息通信研究院、蜜度科技股份有限公司、浪潮云信息技术股份公司、中电信数政科技有限公司、上海梦创双杨数据科技股份有限公司、仪电双杨智能科技（上海）有限公司、北京百度网讯科技有限公司、阿里云计算有限公司、中电万维信息技术有限责任公司、天翼云科技有限公司、联通数字科技有限公司、华为技术有限公司、广电运通集团股份有限公司、浪潮软件科技有限公司、中移（苏州）软件技术有限公司、中电信数智科技有限公司、中电云计算技术有限公司、杭州余杭大数据经营有限公司、腾讯云计算（北京）有限责任公司、北京致远互联软件股份有限公司、广电运通集团股份有限公司、浙江百应科技有限公司、中国铁塔股份有限公司

本文件主要起草人：栗蔚、徐恩庆、张琳琳、王昉、吴佳兴、吴宁、宋佳明、宋光通、王宁、陈尧、薛娇、张宜梅、张龙、刘凤月、姜帆、于希光、张睿智、田丰、王栋梁、张宝玉、李宗倍、秦祎晗、宋建平、徐祯琦、郜成刚、宋汝良、戴鸿轶、苏志伟、苑辰、谷颖慧、王岭钢、任俊龙、马俊国、周建华、钟明康、刘增志、王鹏、崔昊、冯晓蒙、吴至婧、徐天适、张文静、肖思琪、刘刚、樊丁丁、孟于杰、何煦、汤良、王伟印、祁超、王鲁、李存冰、李照川、石园、王珂琛、顾阳、唐紫浩、罗彬彬、王永霞、高新珉

引　言

以人工智能大模型为核心的新一代信息技术，凭借其强大的语义理解、多模态处理及智能决策能力，正逐步成为提升政务效率、优化决策流程、强化城市治理的重要技术支撑。然而，政务大模型在技术标准统一性、场景适配性、安全可控性等方面仍面临挑战。为进一步规范政务大模型的技术研发与应用实践，推动其在政策分析、智能办公、市民服务等场景中的规范化落地，围绕政务大模型的生产能力、通用服务能力、场景服务能力、运营支撑及安全保障五大维度制定此标准，明确模型开发、知识管理、多轮交互、数据安全等核心要求，旨在构建技术统一、安全可信、场景适配的政务大模型能力框架。本标准的实施将为数字政府智能化转型提供技术基准，促进政务领域人工智能技术的健康发展。

政务大模型通用技术与应用支撑能力要求

1 范围

本文件规定了政务大模型通用技术与应用支撑能力要求，包括五方面：一是政务大模型生产，二是政务大模型通用服务能力，三是政务大模型场景服务能力，四是政务大模型运营支撑，五是政务大模型安全保障。

本标准适用于：

a）政务大模型服务提供方评估自身技术与应用能力；

b）政务大模型应用方或管理方对政务大模型服务提供方的服务能力要求；

c）第三方评估政务大模型服务提供方的能力。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件，仅注日期的版本适用于本文件。凡是不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 41867—2022 信息技术 人工智能 术语

GB/T42018—2022 信息技术 人工智能 平台计算资源规范

GB/T 45288.1—2025 人工智能 大模型 第1部分：通用要求

3 术语和定义

GB/T 41867—2022界定的以及下列术语和定义适用于本文件。为了便于使用，以下重复列出了GB/T AAAAA—AAAA中的某些术语和定义。

3.1

模型优化 model optimization

提升模型执行速度，泛化能力，或改善利益相关方所关心的其他特性的方法。

[来源：GB/T 41867—2022，3.2.19]

3.2

强化学习 reinforcement learning；RL

一种通过与环节交互，学习最佳行动序列，使回报最大化的机器学习方法。

[来源：GB/T 41867—2022，3.2.25]

3.3

大模型 large-scale model

大规模深度学习模型 large-scale deep learning model

基于大量数据训练得到,具有复杂计算架构,能处理复杂任务,且具备一定泛化性的深度学习模型。

注：大模型的参数有其功能和模态决定，一般不低于1亿规模。大模型训练使用的数据总量受参数量的影响，达到收敛的大模型的参数量的对数与其训练数据总量的对数成正比。

［来源：GB/T 45288.1—2025，3.1］

3.4

大模型服务 large-scale model service

开发、应用大模型及大模型系统的服务,以及以此为手段提供支持需求方业务活动的服务。

注:常见大模型服务内容包括大模型平台服务、大模型开发定制服务、大模型推理及运营服务.

［来源：GB/T 45288.1—2025，3.2］

3.5

微调 fine-tuning

为提升机器学习模型预测准确性,使用专门领域数据在大模型上继续训练的过程。

注1:专门领域数据一般是特定场景的生产数据或合成数据。

注2:常用的微调方法包括提示词微调、全参微调、参数高效微调等。

［来源:GB/T41867—2022,3.2.31,有修改］

3.6

提示词 prompt

提示语

使用大模型进行微调或下游任务处理时,插入到输入样本中的指令或信息对象。

［来源：GB/T 45288.1—2025，3.5］

3.7

政务大模型 government Large-Scale Model

政务大模型是指基于大规模数据训练和深度学习算法构建，专为政务领域设计的人工智能模型，具备强大的语义分析、信息处理及预测能力，旨在提升政府决策效率和服务质量，优化政务流程，并实现政务服务的智能化与精准化。

4 符号和缩略语

下列符号和缩略语适用于本文件。

API 应用程序接口(Application Programming Interface)

QA 问题与答案(Question-Answering)

BI 商业智能(Business Intelligence)

SSL 安全套接层(Secure Sockets Layer)

TLS 传输层安全性协议(Transport Layer Security)

NeRF 神经辐射场(Neural radiance fields)

CV 计算机视觉(Computer Vision)

OSS 对象存储（Object Storage Service）

3D 三维技术（Three-Dimensional）

NeRF 神经辐射场（Neural radiance fields）

5 政务大模型通用技术与应用支撑能力框架

图 1 政务大模型通用技术与应用支撑能力框架

6 政务大模型生产能力

6.1 政务知识管理

6.1.1 政务知识资源库建设

a) 应支持对知识进行分类管理，帮助用户进行精细化的知识管理和分类，提升查找和使用知识的便利性；
b) 应支持对公共知识属性进行管理，支持用户自定义创建和对已有的属性进行修改、删除等操作，以便于搜索和过滤；
c) 应支持类目知识属性管理能力，支持对知识进行分类管理，用户可以自定义创建类目并将知识分配到相应的类目中，如地理空间、政务政策、法律法规、机构职能、政府公示等。通过类目知识管理，用户可以更好地对知识进行分类，方便搜索和使用；
d) 应支持用户录入和扩展使用政务行业专属知识，并支持多种数据结构的录入、配置和预览，以及灵活的文件上传方式。用户可以将自己的专业知识录入到系统中，并且这些知识能够在政务大模型中被扩展应用到其他相关场景或功能中，以提高模型的效能和准确度。

6.1.2 政务知识资源质量控制

a) 应支持知识自动打上政务标签；
b) 应支持知识库的智能摘要，确保知识的精简融合；
c) 宜设立政务知识资源审核机制，确保知识的准确性和合规性。

6.2 模型开发

6.2.1 数据采集处理

6.2.1.1 数据收集

a) 应支持常见数据源，包括支持线上OSS，线下文件作为数据源接入；
b) 应支持多种数据类型，支持excel、txt、json等多种格式数据导入，以及支持结构化数据、非结构化文本、音视频等多模态数据接入，提供数据去重工具。

6.2.1.2 数据标注

a) 应支持微调语料标注能力，即对已有大规模通用语料库进行精细化标注，以满足特定任务或领域的需求。标注结果应具备一致性和可靠性，遵循相应的标注规范；
b) 应支持对齐语料标注能力，具备将不同来源、不同结构的文本进行整合和对齐的能力，形成一致、规范的文本数据。考虑文本的语法、语义和上下文信息，确保文本对齐和一致性；
c) 应支持QA抽取能力，从原始文本中提取问题与答案相关信息的能力，以构建问答对数据。关注问题的表述清晰度、答案的准确性以及上下文关联性，确保问答对质量；
d) 应确保标注过程的可追溯性，便于后期数据审核和质量控制；
e) 宜支持多人协同标注，提高标注效率。

6.2.1.3 数据增强

a) 应支持 prompt 扩写能力，具备对输入 prompt 进行拓展的能力，通过引入更多的描述性词汇和细节信息，以丰富原始数据的表达形式。这有助于提高模型的泛化能力和表现力；
b) 宜支持对原始数据中的关键词或词组进行近义词替换，从而增加数据的多样性。近义词替换应考虑词义、语法和上下文信息的相似性，以确保替换后的文本在语义上保持一致。

6.2.2 部署推理服务

a) 应具备高性能、稳定可靠的服务器资源，满足大模型的计算和存储需求，提供高速、稳定的网络资源，确保大模型数据传输的实时性和准确性；
b) 应使用稳定、安全的操作系统，具备高效稳定的模型推理引擎，提供完善的数据管理机制，确保数据的完整性、一致性和安全性；
c) 应采用SSL/TLS或其他加密技术，确保数据传输过程中的安全性；
d) 应具备完善的权限管理机制，防止未经授权的访问和操作及完备的日志记录和实时监控机制，确保系统的稳定性和安全性；
e) 应具备自动或手动的故障恢复机制，确保推理服务的高可用性并采用负载均衡技术，均匀分配请求，防止单点故障；
f) 应提供定期更新和升级推理服务相关组件，确保系统的安全性和稳定性；
g) 宜具备根据需求弹性扩展或缩减推理服务资源的能力，支持同时部署和管理多个政务大模型，提高资源利用效率。

6.2.3 模型选择与配置

a) 应支持多种自研通用大模型的选择和配置，以满足政务大模型的需求；
b) 应支持可以根据具体任务选择合适的模型。

6.2.4 模型管理

a) 应支持创建、存储和管理多个自研大模型的模型库；
b) 应支持模型计算性能监控管理能力，实时监测模型的计算资源利用率、响应时间等指标，及时发现并解决性能问题；
c) 应支持政务大模型版本管理的能力，对不同版本的政务大模型进行管理和追踪，方便用户回溯和比较不同版本的性能差异；
d) 应支持对模型库进行权限管理，确保只有授权人员可以访问和操作模型库中的模型；
e) 应支持记录微调意图管理能力，在模型文档中详细描述每个微调模型的目的、功能和适用场景；这有助于其他团队成员了解模型的设计思路，以便在后续维护和优化过程中做出相应的调整；

f） 应具备政务大模型持续改进运营能力，建立一个反馈机制，收集用户和团队成员对微调模型的意见和建议。根据这些反馈，持续改进模型，提高其准确性和实用性。

6.2.5 模型训练与精调

a） 应具备大模型精调、调优功能，可对全量或部分参数更新，并可对调优后的模型进行保存和版本管理；
b） 应支持大模型对齐训练，以确保在不同政务任务和场景间的知识迁移、共享和一致性，从而提高整体模型的性能和效率；
c） 应支持使用有标签的政务相关数据对预训练模型进行微调，以适应具体的政务任务；
d） 应具备容错和恢复机制，以确保训练过程的稳定性和可靠性。当发生错误或意外情况时，平台能够自动恢复并继续训练；
e） 应提供直观的可视化界面，方便用户监控和管理训练过程。用户可以通过界面查看模型的训练曲线、性能指标等信息，并进行相应的调整和优化。

6.2.6 模型优化

a） 应支持数据回流，支持将线上业务修正的结果保存回数据集，经人工筛选、调整以及优化处理后，用于下一次的模型优化训练，持续提升模型输出准确度；
b） 应支持模型压缩，可通过量化、稀疏化等方法在尽量减少精度损失的前提下，降低算力资源占用，提高推理速度；
c） 应支持对政务大模型进行指令微调，使其能够更好地理解和执行用户指令，可通过构建指令对的训练数据集，并使用高效微调、全量微调等方法进行微调，从而提升政务大模型在多种应用场景下的智能交互体验；
d） 应具备灾难性遗忘防治能力，确保模型在持续学习过程中不会忘记先前知识，可通过输出保留先前学到的知识、对比测试、迁移学习测试、长期训练与验证、用户满意度调查或模型自我评估等方法进行验证；
e） 应具备大模型幻觉处理能力，可通过利用外挂知识库等方法去减轻大模型幻觉影响，采用自动化事实检查、人工审核、建立反馈机制或持续更新知识库等方法去验证，从而避免政务大模型在生成回答时出现不准确或不合理的情况；
f） 应支持强化学习的训练方法，以实现模型在政务场景中的智能决策与优化；
g） 宜具备在推理过程优化的处理能力，如优化后的政务大模型在处理用户请求时，应有相比优化前模型更快的作业完成时间；
h） 宜具备识别相同或相似问题并避免重复回答的能力，可通过引入问题去重机制，对输入问题进行预处理，识别重复问题并跳过已回答过的部分，或通过对模型输出的答案进行去重处理，针对性优化实现。

6.3 模型评测

6.3.1 技术评估

a） 应支持由专业人员对模型进行人工评估，以获取更准确和全面的评估结果；
b） 应支持使用自动化工具和算法对模型进行评估，以提高评估效率和一致性；
c） 应支持将多个模型进行对比评估，以选择最佳模型或进行模型融合；
d） 应支持使用通用的评测集对模型进行评估，以验证模型在广泛任务上的性能；
e） 应支持使用政务行业的专门评测集对模型进行评估，以验证模型在特定领域任务上的性能；

f） 应具备模型评估指标管理相应规范；
g） 应支持自定义评估指标和评估流程，满足不同场景下的需求。

6.3.2 可信评估

a） 应具备在不同环境和条件下通过调试信息展示推理过程，获得稳定且可持续的高准确率的能力，如政务大模型对输入样本扰动具有一定的容忍程度；
b） 应具备对于推理和决策过程的解释能力；
c） 应留存完整审计日志，记录模型推理请求来源、数据调用路径、异常行为捕获等信息，支持动态监测与追溯分析；
d） 应确保核心代码架构符合可溯源要求，在脱敏合规前提下允许对算法逻辑、数据处理模块进行代码审计。

6.4 应用服务组装

6.4.1 应用管理

a） 应支持整合与处理来自不同领域、不同层面的知识，从而更好地满足来自多样化场景的需求。将专业领域的知识与通用语言建模能力相结合，为用户提供高质量的内容和服务；
b） 应预置政务场景大模型应用模板，如政策文件、政务问答、政务要素等，模板中需预置行业大模型，大模型需具备政务行业属性，以实现快速构建政务大模型应用能力，同时支持用户自定义调整模板的业务参数，以满足业务的个性化需求；
c） 应支持用户创建应用模板，应用模板支持配置模型、关联模型管理、知识库、应用相关能力的开启关闭等参数，支持在线调试；
d） 应具备大模型应用编排功能，支持用户将多个大模型进行组合和编排，形成复杂的应用流程，满足实际业务需求；
e） 应支持通过可视化交互页面进行效果评测测试；
f） 应支持开放架构，支持第三方工具集成，全流程工具链支持各类AI开发场景，支持外部通过API调用的方式实现智能化应用服务的调用，同时开放接口API需进行鉴权。

6.4.2 生成式应用组件

6.4.2.1 意图组件

意图识别服务能够通过准确理解用户意图，提供更加个性化和灵活的回答和服务，从而提升用户体验。同时，还可以结合垂直领域的知识库和机器学习技术，提供更加精准和高效的意图识别服务。

a） 应具备理解用户目的能力，意图识别服务能够通过对用户输入的文本进行分析，理解用户的意图，从而为后续的对话处理提供明确的方向；
b） 应具备自然语言交互的能力，意图识别服务可以支持自然语言交互，用户可以使用自然语言进行提问和查询，无需特定格式或关键词，使得对话更加自然和便捷；
c） 应具备多轮对话支持的能力，意图识别服务可以支持多轮对话，可以在对话过程中动态地调整意图识别结果，从而更好地引导对话过程；
d） 应具备强大的后端知识库，意图识别服务通常与强大的后端知识库相结合，可以提供更加全面和准确的信息和服务；
e） 宜具备自我学习和优化的能力，意图识别服务可以通过机器学习和自然语言处理技术进行自我学习和优化，不断提高对话系统的效率和准确性。

6.4.2.2 检索组件

检索组件允许用户高效地查询和获取文档中的关键信息，提升信息获取的效率，进而提升工作和决策的效能。

a) 应支持对表格、文本数据进行接入、上传，支持对详情进行预览查看；
b) 应支持配置检索库生成任务过程的配置，支持数据与处理算法、策略的配置和自定义；
c) 应支持对政策文件、公文等政务主流数据的专属检索策略；
d) 应支持检索库的创建、配置、查询；
e) 应支持检索库的全量更新、增量更新、定时更新、手动更新；
f) 应支持检索库以API的形式进行调用。

6.4.2.3 提示（prompt）组件

提示(prompt)组件在大模型服务中可以通过自然语言的交互形式引导模型完成指定的任务和目标，同时起到引导用户输入、约束输入范围、提高用户体验和优化模型预测的作用。

a) 应具备引导用户输入的能力，prompt组件可以引导用户输入，例如提示用户输入某个关键词或问题，以便引导后续的对话处理；
b) 应具备提高用户体验的能力，prompt组件可以提供自然语言交互和多轮对话支持，使用户可以更加自然和便捷地进行对话，提高用户体验；
c) 应具备优化模型预测的能力，prompt组件还可以优化模型预测结果，通过提供特定的关键词或问题，引导模型更加准确地预测用户意图和提供相关回答；
d) 应支持prompt工程优化，对Prompt话术模板进行内容上和结构上的优化，便于获得更符合期望的大模型推理质量和结果；
e) 宜具备约束输入范围的能力，prompt组件可以约束用户输入的范围，例如限制用户只能输入特定类型的问题或查询关键词。

6.4.2.4 会话日志组件

会话日志组件记录了用户与AI系统的对话历史和交互信息。这些信息对于系统开发者和管理员来说非常重要，因为它们可以帮助了解用户的需求和行为，从而优化AI系统的性能和用户体验。

会话日志组件应包括以下内容：

a) 应包括对话历史记录，会话日志记录了用户与AI系统中被允许收集的对话历史，包括文本输入、语音交互、表情符号等。这些对话记录可以展示用户的需求和问题，以及AI系统的回答和反馈；
b) 应包括交互信息，会话日志还记录了用户与AI系统的交互信息，包括用户输入的时间、输入方式、AI系统的响应时间、回答内容等。这些信息可以帮助开发者了解用户与AI系统的交互方式和效果，从而优化系统的性能和用户体验；
c) 应包括意图识别结果，会话日志记录了AI系统对用户输入的意图识别结果，包括识别到的关键词、话题、任务等。这些信息可以帮助开发者了解用户的需求和行为模式，从而优化AI系统的意图识别算法和性能；
d) 宜包括情感分析结果，会话日志还可以记录AI系统对用户输入的情感分析结果，包括用户的情绪状态、情感倾向等。这些信息可以帮助开发者了解用户对AI系统的情感反馈，从而优化系统的情感分析算法和性能；

6.4.2.5 安全组件

安全和幻觉控制模块是生成式AI系统的重要组成部分，它们可以提供强大的安全防护、幻觉控制、可解释性和透明度等功能，并具有灵活性和可扩展性。

a) 应具备安全防护的能力，这些模块可以保护生成式AI系统的安全，防止恶意攻击和数据泄露。它们提供了数据加密、访问控制、安全审计等功能，以确保模型和数据的机密性和完整性；
b) 应具备幻觉控制的能力，幻觉控制是指通过技术手段实现对AI系统的控制和操纵。这些模块可以帮助开发者更好地理解和控制生成式AI系统的行为。例如，开发者可以通过幻觉控制来引导AI系统的决策过程，以确保系统能够按照预期的方式运行；
c) 应具备灵活性和可扩展性的能力，这些模块通常具有灵活性和可扩展性，可以适应不同的应用场景和需求。开发者可以根据具体的应用场景来选择适当的安全和幻觉控制模块，并对其进行定制和扩展；
d) 宜具备可解释性和透明度的能力，这些模块可以提高生成式AI系统的可解释性和透明度，使开发者能够更好地了解AI系统的内部工作原理和决策过程。这样可以帮助开发者更好地理解系统的行为，并发现潜在的安全风险。

6.4.3 插件库管理

在大语言模型中，插件是一种扩展功能，可以提供更加个性化和高效的语言处理体验。插件可以增强大语言模型的能力，使其能够访问新的、私人的或具体的，不包含在训练数据中的信息，还可以使大语言模型代为执行安全、受限的操作，从而提高整个系统的实用性。

6.4.3.1 智能知识文档插件

智能知识文档插件支持外部文档接入到系统中作为知识补充，并优先参考文档知识内容。

a) 应支持文档知识库的创建、删除等功能，以满足用户对知识文档的分库管理；
b) 应支持doc、docx、wps等若干种文档格式，同时支持单文档或者批量文档的上传、下载、删除、查询等功能，以方便用户修改知识文档库；
c) 应具备文档理解的能力,能够自动解析和提取政务文档中的关键信息,包括但不限于政策法规、通知公告、办事指南等。它应该具备对文档进行自动化的结构化分析能力，将文档中的重要信息提取出来，如主题、观点、论据等，帮助用户更好地掌握文档要点，以便于后续的问答处理；
d) 应具备自动摘要的能力，能够自动生成文档的摘要，方便用户快速了解文档内容；
e) 应具备文档分析的能力，能够分析文档的结构、语言风格、论述逻辑等，帮助用户更好地理解文档的写作特点和优缺点；
f) 宜具备内容建议的能力,通过分析文档内容,可以提供与文档相关联的其他资料或资源的建议，以扩展用户的阅读范围；
g) 宜具备跨语言支持的能力,支持多种语言,可以为不同语言的用户提供高质量的文档理解服务。

6.4.3.2 智能数据分析插件

智能数据分析插件帮助大模型应用针对数据表进行处理以及数据结果召回,用户需要新建对应的数据分组并绑定关联相关的数据库数据表，关联好的应用即可使用智能数据分析能力。

a) 应支持自定义数据源，支持接入多类型、多来源的数据，同时对数据进行预处理；
b) 应具备注册外部数据分析服务及其语义描述的能力；
c) 应具备自动调用外部相关数据分析服务并完成对应数据分析操作的推理行动能力；
d) 应能够从数据源中自动识别和提取与问题相关的信息,根据从数据源中获取到的信息进行数据分析与推理，以满足用户需求；

e) 应支持将分析结果以直观的方式呈现给用户，以便用户能够更轻松地理解数据。数据可视化可以使用各种图表、图像等，将复杂的数据转化为易于理解的形式。

7 政务大模型运营

7.1 业务运营

a) 应具备收集用户反馈和业务需求能力，从而对业务进行相应的调整和处理，以支持模型的持续优化和改进；
b) 应支持业务端到端专家评测所需的测试集管理功能，支持单论及多轮评测集的构建与查询；
c) 应支持业务评测任务的创建、分配、回收等管理能力，并支持基于大模型为预设任务生成答案；
d) 应支持业务问答评测，包括单论、多轮预设问答与自由问答能力；
e) 宜支持知识依据反馈，具备知识依据过期或错误标记的能力。

7.2 平台运营

a) 应支持多账户与多角色管理能力，支持大模型以及政务AI原生应用的统一权限管理；
b) 应支持基于角色的工作流设计，实现业务流程的自动化流转，应提供任务协同功能，支持不同角色间的在线协作，并监控和记录业务流程的执行状态，确保流程的合规性和高效性；
c) 应具备平台运维能力，包括服务器、存储、网络等基础设施的部署、配置、监控和故障排除等，确保平台的稳定性和可用性；
d) 宜支持各类接入服务的运营管控能力，支持发布、审批、申请、回收等资产管理控制功能。
e) 宜支持查询平台各政务大模型服务以及各业务需求方的历史调用量；
f) 宜具备平台扩展能力，能够根据业务需求进行平台的扩展和升级，包括增加计算资源、存储容量、用户规模等，以满足不断增长的模型训练和应用需求。

7.3 数据运营

应设立专门的敏感词库，包括法律法规禁止使用的词汇、不当言论等，定期更新、维护，确保库内词汇的准确性和实时性。

a) 宜具备对大量数据的存储、检索和管理的能力。这包括提供高效的数据库系统，支持数据的备份和恢复，以及提供数据的安全保护措施；
b) 宜支持对数据服务、数据表、文件三种数据资产申请与资产运营管理；
c) 宜支持平台内数据资产态势概览，包括数据变化趋势、申请热度，来源分布分析等。

7.4 服务运营

a) 应具备统一的服务接口规范，支持服务的动态编排组合，同时提供可复用的服务封装能力；
b) 应具备管理一体化能力，能够实现对推理服务的管理和调度，包括服务的创建、分配、监控和优化等；
c) 应具备用户支持与服务能力，提供用户支持和服务，包括技术咨询、培训服务、使用指南、在线测试、故障处理等，帮助用户更好地使用推理服务；
d) 应具备服务运营数据分析能力，能够对推理服务的运营数据进行收集、分析和挖掘，为优化服务提供数据支持；
e) 应支持政务大模型的调用量统计分析，不限于用量排行、调用成功率、使用率等指标；
f) 应支持政务大模型服务的性能统计分析，包括资源占用、服务稳定性统计等指标；

g) 应支持政务大模型以及政务AI原生应用的线上使用情况统计与分析能力。

8 政务大模型通用服务能力

8.1 交互式问答

a) 应具备准确理解和回答用户问题的能力，包括对问题的语义、语境和意图的深入解析，同时给出准确、简洁且符合事实的回答，避免产生歧义或误解；
b) 应具备对话管理的能力，能理解和遵循用户的意图和需求，进行连续、有逻辑的对话，以提供连贯、准确的服务；
c) 应支持对问答业务策略进行配置，支持根据场景需求，提供可自定义配置的开场话术和推荐问题，来确保面向政务各类业务领域能快速拓展；
d) 应支持根据所依赖的政务知识进行回答，在交互过程中能有效地从知识库中获取相关信息，并能在回答中进行引用和解释；
e) 应具备对用户问题和大模型回答的内容审核能力，包括敏感信息过滤、恶意言论识别、违法信息筛查、回答内容的审核等，以提供准确、可靠、安全的交互式问答服务；
f) 宜支持根据用户的问答内容，给出推荐问题，提升用户提问效率。

8.2 生成式 BI

a) 应支持通过单属性、多属性等简单条件对于对数据库表进行检索查询。
b) 应支持通过文本模糊匹配、去见判断、多条件逻辑计算等复杂条件对于对数据库表进行检索查询。
c) 应支持基于列名进行处理后计算，来进行检索查询。
d) 应支持基于用户自然语言Query找到对应数据表。
e) 应支持基于对话的多类型数据库查询能力，基于自然语言的交互形式进行底层数据的搜索及回复。
f) 应支持基于对话的多类型图表生成能力，可基于用户问题的理解，构建相应的可视化图表，如柱图、饼图、线图等。
g) 应支持数据分析及洞察能力，可基于用户问题，分析数据中的模式和趋势，提供相应的洞察及建议。
h) 宜支持人机协作能力，支持用户对大模型产生的中间结果，如数据库查询等，进行查看和修改，以确保获取数据的正确性。

8.3 多轮引导反问

a) 对于普通对话场景，应提供多轮闲聊问答能力。
b) 应提供多轮常识类问答以及对问答中意图和对上文问题的修改能力。
c) 应支持以职位为基础的问答，如以具体角色如客服人员的口吻对问题进行回答。
d) 应支持基于业务流程对于用户的对话内容进行意图识别，意图不明确时支持澄清范围。
e) 应支持基于流程对关键信息进行提取，对于未获取的关键引导信息进行澄清反问。
f) 应支持对于某事项的必要条件进行澄清，不确定时进行澄清反问、追问。
g) 应具备对问题边界识别的能力，能自动识别用户提问是否超出当前大模型的能力覆盖范围，对于超出能力边界的问题，能够以适当的方式拒绝回答，如可选择用兜底拒绝话术进行回答，或拒绝答复，并用标准术语答复，避免大模型在能力边界外产生不合理的回复。
h) 宜支持对用户问题进行相似问题推荐的能力。

8.4 数据定期更新能力

a) 应支持通过外挂知识库进行数据的人工或定期更新。

8.5 文案生成

a) 应具备遵循明确的语言规范和标准的能力，生成的文案应语法正确、用词准确、表达清晰，不出现错别字、语法错误和语义歧义。
b) 应具备对输入信息进行准确理解和处理的能力，以保证生成文案的准确性。这包括但不限于对输入信息的语义理解、实体识别、关系抽取等能力的掌握，以及能够根据这些信息生成准确、完整的文案。
c) 应具备语言表达的流畅性，应掌握不同文体的表达方式，生成的文案应语句通顺、逻辑清晰、符合语言习惯，避免出现语病和表达不畅的情况。
d) 应具备多语言支持能力，能够处理不同语言的输入信息并生成相应语言的文案。具备对多种语言的处理能力和掌握多种语言的语言资源库。
e) 应支持标签生成、摘要生成、文案生成等能力，支持用户自定义文案生成的要求，如数量、字数、主题等。
f) 应预置政务场景文案生成能力，如：公文生成、公文摘要、政务标签、政务新闻稿等。

8.6 计算机视觉（CV）

a) 宜支持物体检测，基于CV大模型的物体检测工作流，具备自动化抽取和识别图像中特定物体（如车辆、行人等）的能力。通过模型的自动调参和抽取，可以适应不同物体检测场景，从而提高政务管理的效率和准确性。
b) 宜支持对人体关键点检测和姿态估计，基于CV大模型的姿态估计工作流，以便在视频监控、人像识别等场景中准确识别个体姿态，为政务管理提供便捷手段，助力安全监控和人员管理。
c) 宜支持视频分类，基于CV大模型的视频分类工作流，实现对视频的高精度分类，如公共安全、城市管理、民生服务等类别，以便快速筛选和定位关键视频信息，提高政务处理效率。
d) 宜支持图像分类，基于CV大模型的图像分类工作流，实现对图像的高精度分类，如文档识别、车牌识别等，以便快速识别和处理政务场景中的图像信息。
e) 宜支持异常检测，基于CV大模型的异常检测工作流，非正常即异常，用于如识别政务数据中的异常值或异常模式，以便及时发现潜在问题和风险。
f) 宜支持目标跟踪，基于CV大模型的目标跟踪工作流，实现对物体的位置检测和跟踪，如车辆跟踪、人员定位等，为政务管理提供实时监控能力。
g) 宜支持语义分割，基于CV大模型的语义分割工作流，根据图像语义将不同区域分割成不同类别，如道路、建筑、绿化带等，有助于政务场景下的城市规划、环境治理等应用。
h) 宜支持实例分割，基于CV大模型的实例分割工作流，对物体的位置进行检测并且识别出物体的轮廓信息，如车辆轮廓、建筑轮廓等，有助于政务场景中的精确识别和定位。
i) 宜支持图像增强，基于CV大模型的图像增强工作流，实现对图像的画质及目标成像增强，如画质模糊、过曝、失真等，有助于政务场景中CV算法的目标检测和识别。
j) 宜支持开放语言的目标检测和分割，基于CV大模型的工作流，根据开放性地输入目标名称，对图像进行检测和分割，如道路轮廓、车辆类别等，有助于政务场景中的精确识别和定位。
k) 宜支持深度估计，基于CV大模型的深度估计工作流，实现对图像中的目标进行距离估计，如人员或车辆距离摄像机的实际距离等，有助于政务场景中目标在实际环境中的精确定位。

8.7 多模态处理

多模态大模型是指模型中涉及到了两种及其以上的模态信息。目前，大多数的多模态模型只涉及了两种模态之间的建模，例如文本-图像，文本-语音等。根据政务场景中常见的数据类型，模型应支持图像、文本、3D 和语音等模态的处理。

宜支持以下能力：

能力项	场景	场景描述
图像生成	以图生图	宜具备自动理解输入图像的语义、构图和风格，和文本输入的图像属性要求，进行图像重构。
	以文生图	通过自然语言生成对应语义的图像。宜支持泛化的语言概念和丰富的风格，支持大分辨率（1080P 以上）图像生成。
	可控生图	宜支持根据线稿、深度图、人物姿势等控制信号生成相对于的图像。
	概念植入	宜支持根据少量图像的快速训练，拷贝相关视觉概念并根据指定概念生图。
	图像编辑	宜具备对图像修复、水印、对象修改、空洞补全、根据姿态或者线稿等信号进行图像生成。
图像视频理解	图像视频描述	宜支持基于给定图像、视频，结合文本描述要求，进行对象、属性、情景、行为等的客观描述和深层次语义理解，给出城市治理业务中的对象或事件的标签、短句或者长句描述。
	图像视频问答	宜支持基于给定的城市治理业务中的对象或事件的图像、视频，进行对象，属性、情景和行为识别和理解，结合问题，进行精准答复。
	文本图像检索	宜支持基于给定的城市治理业务中的对象或事件图像、视频库，结合给定的文本，进行对应的 embedding（向量化），对图像、视频进行检索，找到业务最相关内容
语音识别	语音识别	宜支持将语音转化为文字，可用于会议记录、智能客服、文件整理、电话服务等，提高工作效率，减少录入错误。
	语音理解	宜支持对语音内容和语气语调进行综合分析，支持服务评估、话务质检、报告生成等政务场景中的语音理解任务。
	语音生成	宜支持将文字转化为语音，可根据服务对象进行个性化语音生成，包括但不限于方言、情绪、语速等方面，可用于自动电话服务、虚拟助手和语音导航等，提升公共服务效率。

3D 生成	3D 生成-3D 空间新视角生成	根据原始视频+雷达点云（可选），基于 NeRF 生成大空间新视角视频，比传统 3D 重建+人工修模+渲染，效率提高 2+倍，视觉真实度大于传统渲染方案；可同时生成视频、点云、以及对应的 label 信息；
	3D 编辑-3D 空间内容增删	根据原始视频+雷达点云（可选），基于 NeRF 删除、增加物体，物体支持静态、动态物体，比如静态障碍物和移动的车辆，比传统 3D 重建+专业 3D 软件人工设计+渲染，效率提高 2+倍，视觉真实度大于传统 3D 软件方案；可同时生成视频、点云、以及对应的 label 信息；
	3D 可控生成-设计图对齐的 3D 漫游生成	基于轻量化全景相机（可选），在建筑项目采集视频，云端自动对全景视频进行 3D 建模、语义对齐，生成和设计图对齐的 3D 全景漫游、跨日期对比 3D 全景漫游。
	3D 动作生成	根据自然语言描述（可选），自动生成人物模型的运动序列，用于驱动 3D 数字人。

9 政务大模型安全保障能力

9.1 数据安全

应具备制定全面、科学的数据去毒策略、去毒效果评估和去毒过程审计与追溯的能力，确保有毒数据在训练过程中不被使用。有毒数据包括但不限于：事实错误、涉恐、涉黄的敏感数据、不符合伦理道德的数据等。

a) 应具备数据来源可追溯性的能力，确保数据来源有据可依，避免使用恶意或有意篡改的数据。
b) 宜具备评估政务大模型训练、部署、推理时数据的完整性和一致性，防止数据被篡改或恶意注入潜在的安全漏洞。
c) 宜支持敏感数据处理能力，对敏感数据不应使用弱加密，不应采用明文形式存储或网络传输敏感数据。
d) 宜具备数据清洗能力，涉及到用户隐私的数据需在获得许可后进行收集，并进行必要的数据清洗。
e) 各地区政务大模型的数据管理宜支持符合各地区相关政策管理、条例，确保数据的合法使用和管理，可通过建立合规性管理制度，明确责任和权限，加强对数据的监管和审计。同时，应定期进行合规性检查和培训，确保员工了解并遵守相关法规要求。
f) 宜提供数据安全管控及审计能力，防止数据泄露、篡改等安全风险，可采用防火墙、入侵检测、安全审计等技术手段，确保数据安全。

9.2 模型安全

9.2.1 模型训练安全

a) 应支持模型训练安全，包括但不限于提供多种手段多种手段防范外部恶意攻击、 面对恶意攻击（如数据投毒、后门攻击）的应具备相应的防御性能力、具备对政务大模型实时监控与预警的能力、对政务大模型恢复与应急处理的能力等。

9.2.2 模型推理安全

a) 政务大模型在推理阶段宜具备相应安全能力，包括但不限于对输入数据中的对抗样本进行检测的能力、采用鲁棒性优化技术、对对抗攻击事件的日志记录和审计功能等。

9.2.3 模型文件安全

a) 应具备对模型文件提供安全能力，包括但不限于采用加密技术对模型参数和训练数据进行保护、实施严格的访问权限管理、对政务大模型服务的访问日志进行审计和监控、政务大模型服务中嵌入隐蔽的水印信息、明确政务大模型服务的版权归属等。

9.3 话术安全

a) 应具备有检查话术安全的能力，包括禁用词过滤、敏感词过滤、敏感内容图片过滤等。
b) 应遵守伦理原则，包括尊重个人的自主权利、保护弱势群体的利益、确保公正和公平等的能力。
c) 应具备有隐私保护的能力，包括使用隐私识别技术、隐私脱敏技术、隐私伪装技术等。
d) 应具备一定的鲁棒性，对人为诱导情境下的偏见、毒害、攻击、不确定性具备鲁棒控制能力。

9.4 服务安全

9.4.1 安全组织与机制

a) 应具备专门的安全管理部门和人员，明确相应的岗位职责。
b) 宜具备明确的安全测评框架和安全测评机制，对模型自身、数据、平台等维度实现全方位测评。

9.4.2 访问控制

a) 应对授权访问的内容严格访问控制，不应有超出授权范围的访问，当第三方访问被保护的用户数据时，应先获得用户许可或同意。
b) 不应拦截或存留用户敏感或隐私信息，如用户支付密码等。
c) 应支持细粒度权限管控能力。
d) 应支持安全审计，能够实时监控、记录各类用户的登录和操作行为，并可以进行集中分析并生成审计报表。

9.4.3 可靠性

a) 应支持政务大模型服务多方调用的统一流量管控能力。
b) 应支持政务大模型服务历史日志监控与持久化能力。
c) 应支持政务大模型服务运行状态的监控能力。
d) 宜支持业务激增时保证大模型稳定服务的按需扩缩容能力。

10 政务大模型场景服务能力

10.1 多角色适应性（服务对象）

a) 宜具备为多种角色提供服务的综合能力，至少能够满足其中一种角色，如市民、企业、政府委办局人员、客服人员、巡查员等角色的需求，确保为各角色提供准确、高效的服务，以保证政务大模型在多元化用户场景下发挥最大价值。

10.2 政务场景应用能力

10.2.1 政务服务类

a) 针对市民、企业、客服人员等服务对象，宜提供政务咨询服务，具备自动化处理和智能分析咨询问题之能力，为服务对象提供准确、及时的答复，并涵盖意图识别、交互式对话、开放搜索、数据实时更新、条件补齐、安全回答等全方位能力，以确保高效便捷的政务服务体验，更好地满足各类服务对象的需求。
b) 宜支持交互式智能政务办事能力，提供24小时在线的交互式政务服务，政府和群众可以通过语音、文字等方式与大模型进行交流，实现业务咨询、办理、查询等功能的自助服务，涵盖办事指南能力、智能引导、实时更新指南内容、在线预约办理、智能预审、在线支付功能等能力，帮助市民更好地了解和办理各类事务。
c) 宜支持自动接听市民热线电话，并具备智能分类、解答市民问题的能力；同时提供政务服务便民热线工单实时填报功能，涵盖语音识别与转换、自然语言理解、工单分类与格式化、实时数据处理、数据隐私保护、可视化展示、智能提醒及全流程对话交互式导办服务等能力，实现语音/文本多轮交互式问答对话，以全方位满足市民需求，提升政务服务体验。
d) 宜支持目标客商分析的能力，具备企业检索能力，支持快速检索企业，并提供详细的企业信息。

10.2.2 城市治理类

a) 宜支持对城市的公共事务进行全面管理和优化，通过对大量数据的分析和挖掘，政府可以及时发现和解决各类公共事务问题，提升城市管理的水平和效果。例如，通过分析交通数据和人流数据，政府可以合理规划道路交通，优化公共交通线路，提升交通运行效率。
b) 宜支持市民人员和城市巡查员进行事件上报能力，涵盖图文对话服务、图片语言描述生成、图片要素提取和图文比对等能力，以确保上报事件的准确性和全面性。
c) 宜支持通过城市感知设备获取的视频等数据，运用政务大模型智能识别技术进行分析，以实现事件上报功能，可包括但不限于以下能力：政务服务、城市治理领域的识别算法能力，以及自动提交规则的配置、相似违规事件的识别和重复事件的辅助合并功能。根据事件名称、违规图片、上报时间、上报摄像头、切片视频等信息，自动上报事件，由人工操作员进行误报处理、确认重复或进行上报。事件信息和报警信息将被精准推送给相关单位、部门及人员，从而实现事件智能上报。
d) 宜支持事件派发或事件分拨能力，涵盖事件信息抽取、事件文本摘要、事件文本分类、部门智能推荐、相似事件识别等能力，从而有助于实现事件的高效响应、精准处理与资源的优化配置，提升城市治理的智能化水平。
e) 宜支持事件分析能力，具备智能分析链路生成、在线对话分析、报告撰写服务、事项智能聚合、情感智能分析、报告指标定义与采集任务管理、报表模板管理与生成等功能，以实现对事件的多维度查询、分析摘要生成、可视化建议及历史报告识别，提高事件处理的效率和质量。通过这些能力，能够实现对事件数据的深度挖掘和分析，为决策者提供有力支持，提升整体业务运营水平。

10.2.3 政策分析类

a) 政务大模型通过对政策文件、案例等数据的深度学习，宜能够自动识别符合特定条件的企业或个人，提供精准的政策服务，提高政策落实的准确性和效率。

b) 宜具备政策解读提供对政策文件进行深入解读，为用户提供详细的政策解读和解读报告，提高政务服务的透明度和公信力，增强用户对政务服务的信任和认可。例如，针对某项税收政策，可以生成详细的政策解读报告，帮助用户理解和掌握政策内容。
c) 宜具备政策生成与推演能力，涵盖政企匹配、政策调优建议和文案生成等能力，实现企业画像与政策画像精准匹配，提供符合用户意图的政策调整建议，自动生成政务文本，提升政策制定与执行的精准性和效率，促进政务服务的智能化和便捷化。
d) 政务服务大模型宜对政务数据进行全面深入的分析和挖掘，为政府决策提供科学依据。通过对政务数据的分析，政府可以了解市民的需求和诉求，及时调整政策和服务，提升政府决策的准确性和针对性。
e) 政务大模型宜具备强大的信息检索能力，能够快速、准确地搜索到所需信息，帮助用户更好地了解政策法规和相关业务。

10.2.4 政府办公类

a) 宜支持智能公文写作场景能力，涵盖智能写作、模板智能推荐、内容自动校对、语言智能润色、格式自动调整、用词智能建议以及实时更新等功能，以提升公文写作的效率、准确性规范性和美观性，以确保公文能够有效地传达信息，并符合组织或政府的规范和要求，实现便捷、高效的公文写作体验。
b) 宜支持智能会议场景能力，包括但不限于智能拟制会议方案、预定会议时间和地点、会议资料一键转入会议议题、智能记录会议内容，自动提取会议纪要并与督办、公文、重点工作任务、档案等系统协同对接。
c) 宜支持智能办事场景能力，包括但不限于内跑事项的便捷咨询办理、智能辅助审批、实现多维度考核指标动态生成等能力。

10.2.5 泛政务场景类

a) 宜支持城市经济运行决策分析场景能力，包括但不限于智能问答、指标抽取、图表生成等功能，以提供高效精准的数据分析结果，并通过生成式交互能力，围绕领导视角，助力管理者做出科学决策，实现经济与产业发展管理的综合指标提炼和信息准确表述。
b) 宜支持交通数字专家场景能力，提供知识问答能力，政务大模型能够基于特定上传的文档，自动从给定文档特定领域的知识库或语料库中查找并生成答案。化身“交通数字专家”，为地铁和道路交通运营提供全面、高效的专业知识搜索问答和初级分析研判能力，实现智能客服、智能运维以及应急指挥等多种功能应用。
c) 宜支持档案管理场景能力，包括但不限于知识型搜索能力、对档案知识智能检索的再标注等能力。
d) 宜支持法务数字助手场景能力，包括但不限于仿写型法律意见书生成、文本纠错、安全可靠等能力。
e) 宜支持智慧农业种植助手能力，包括支持对农业产品进行判断、提供图片识别和文本搜索图片等能力，实现对农业种植全方面领域知识的智能检索，通过人机交互解决种植者农业生产理论知识不足、种植管理经验不足等问题。