

基于大模型的政务咨询系统技术要求与评估方法

编制说明

1、标准“范围”的内容

本文件面向以大规模训练模型为技术底座，能够提供智能问答、政务咨询、知识搜索等功能的基于大模型的政务咨询系统。

本文件规定了基于大模型的政务咨询系统的功能、性能要求和评估方法，主要包括大模型基础能力、政务咨询业务能力、系统安全应用能力及指标评估方法四个部分。

本文件适用于基于大模型的政务咨询系统及同类产品的研发、评估和验收等工作。

2、工作简况

2.1 任务来源

目前，国内以拓尔思、讯飞、用友等为代表的技术研究企业，先后推出了政务大模型产品服务，北京市发布了全国首个政务服务大模型场景需求，涉及政务服务管理、中小企业服务等数十个项目，广州市拟成立实验室研究大模型在政务咨询方面的应用，探索大模型驱动城市高效管理的新模式，越来越多的城市和地区开始关注和重视政务大模型的发展和应用。然而，在政务大模型领域的标准缺失，特别是针对基于大模型的政务咨询系统的标准亟需建立。

本标准将对基于大模型的政务咨询系统的大模型基础能力、系统功能、服务性能、系统安全可用方面提出指标和评估方法，以规范基于大模型的政务咨询系统及同类产品的研发、评估和验收等工作。

2.2 立项阶段

本标准于2023年11月在广州数据集团的编制启动会上通过立项，

并成立标准编写组。会议邀请了广州市数字政府运营中心、中国信息通信研究院等单位有关代表参加。

2.3 起草阶段

标准编写组为保证标准的内容符合国内行业发展的技术需求，客观地提出合理、适用的技术指标，编制前期起草组成员对可以提供基于大模型的政务系统的厂商、企业进行了相关调研。

标准在编制过程中起草组内部组织召开了多次的技术讨论会，在标准编写过程中征求国内厂家以及用户单位的意见，形成了征求意见稿。

2.4 起草单位和起草人

本文件由：广州数据集团有限公司、广州市数字政府运营中心、中国信息通信研究院等联合起草。

本文件主要起草人：。

3、标准编制原则和确定标准主要内容(如技术指标、参数、公式、性能要求、试验方法、检验规则等)的依据(包括试验、统计数据)

标准编制的意义：有利于规范基于大模型的政务咨询系统的研发、评估和测试验收，促进政务大模型的技术发展和行业应用推广。

标准制依据 GB/T 1.1-2020 《标准化工作准则 第一部分：标准的结构和编写》、GB/T 35282-2023 《信息安全技术 电子政务移动办公系统安全技术规范》、GB/T 31506-2022 《信息安全技术 政务网站系统安全指南的规定编制》的规定编制。

编写过程中贯彻国家关于积极采用国际标准的政策，密切结合我

国国情，做到技术先进合理、使用方便、切实可行。

本标准的主要内容包括：规定了基于大模型的政务咨询系统的功能、性能要求和评估方法，主要包括大模型基础能力、政务咨询业务能力、系统安全应用能力及指标评估方法四个部分。

4、主要试验(或验证)的分析、综述报告

本标准起草过程中，对多家对可以提供基于大模型的政务系统的厂商、政务系统客户进行了细致的调研，经过多轮讨论，不断调整和丰富能力要求和规范指标体系，最终形成本标准。

5、标准在起草过程中遇到的问题及解决办法；

在本标准的修订过程中，无重大分歧意见和技术问题。

6、与国外标准的关系：包括：采用国际标准和国外先进标准的程度，与国外标准主要技术内容的差异

本标准没有完全对应的国际标准，且未采用国际标准或国外先进标准。

7、修订标准时，说明与标准前一版本的重大技术变化，并列所涉及的新、旧版本的有关条款(可引用标准前言的内容)；废止/代替现行有关标准的建议

本标准与其他标准没有关联关系。

8、说明标准与其他标准或文件的关系(可引用标准前言的内容)，特别是与有关的现行法律、法规和强制性国家标准的关系

无。

9、标准作为强制性标准或推荐性标准的建议

建议本标准作为推荐性团体标准。

10、贯彻国家标准的要求和措施建议(包括组织措施、技术措施、过渡办法等内容)；标准发布后，对国内外业界可能产生的影响

无。

11、标准是否涉及知识产权的情况说明；如标准中含有自主知识产权，说明产品研发程度、产业化基础及进程

本标准未涉及。

12、其他应予说明的事项

本标准未涉及。
