

# 《卫生健康领域生成式人工智能应用指南》 (送审稿) 编制说明

## 一、项目背景

生成式人工智能 (Generative Artificial Intelligence, GAI) 基于深度学习技术, 通过训练数据生成文本、图片、音频、视频等内容的算法模型及相关技术。也被称为大语言模型。随着 ChatGPT 在 2022 年 11 月问世, 以 ChatGPT、DeepSeek、Google Bard、Sora、通义千问、文心一言等生成式人工智能技术井喷式发展, 引发公众的强烈关注。生成式人工智能对人类发展是一个巨大的机遇, 但也可能造成伤害和偏见。

2024 年 1 月, 世界卫生组织为生成式人工智能的医疗应用发布伦理指南 (Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models), 并督促在生成式人工智能医疗应用制定标准的问题上, 各国政府负有首要责任。多个国际学术期刊、团体建议对健康领域工作及科研人员, 要建立使用规范, 甚至是使用执照制度。2023 年发布的《大模型人工智能技术在医疗服务领域应用的专家共识》也指出要加强医务人员对模型使用的监管和指导, 建立合理责任认定规则。

世界卫生组织于 2021 年发布了卫生领域人工智能的伦理和治理指南, 并于 2024 年 1 月份发布了针对生成式人工智能的伦理和管理指导文件 (Ethics and governance of

artificial intelligence for health: Guidance on large multi-modal models)，为原则性指导，指导文件列出了供政府、技术公司和卫生保健机构考虑的 40 多项建议，以政府和研发公司为主，针对使用的卫生健康机构建议数量有限仅有 5 条左右。

联合国教科文组织于 2023 年 9 月发布首份全球指南《生成式人工智能教育与研究应用指南》，力促对生成式人工智能在教育中的运用实施管制，并帮助政策制定者和教师最大限度地利用人工智能的潜力，以服务于学习者的根本利益。

2023 年 7 月，日本文部科学省初等中等教育局发布首份《初等中等教育阶段生成式人工智能利用暂行指南》，为生成式人工智能在教育现场的应用指明方向。

国内生成式人工智能的标准，有互联网医疗健康产业联盟的《医疗健康行业大模型应用技术要求》标准，规范内容为模型工具的技术要求，如大模型应具备的功能、任务支持度等。中国音像与数字出版协会团体标准《出版业生成式人工智能技术应用指南》。2025 年，《人工智能 大模型 第 1 部分：通用要求》（GB/T 45288.1-2025），《人工智能 大模型 第 2 部分：评测指标与方法》（GB/T 45288.2-2025），《人工智能 大模型 第 3 部分：服务能力成熟度评估》（GB/T 45288.3-2025）发布，指导大模型的设计、开发、部署、应用。

本项目拟制定卫生健康领域应用生成式人工智能的原则和框架，从启动、评估、分类、部署、运维等方面，对卫

生健康机构进行应用方面的指导。制定国际标准能确保这些技术在全球范围内得到安全、有效地应用，同时保障患者权益和数据安全。使其安全高效利用生成式人工智能工具，缓解医疗压力，提升医疗质量，同时减少使用过程中产生的信息安全及隐私风险，保护患者。同时，响应国际组织及专家呼吁，发挥人工智能在医疗领域的作用，推进人工智能与卫生健康融合高质量发展。

## 二、工作简况

深圳市地方标准《卫生健康领域生成式人工智能应用指南》由深圳市卫生健康发展研究和数据管理中心、深圳市卫生健康委员会、深圳市前海蛇口自贸区医院、深圳市福田区妇幼保健院共同起草。

### （一）标准立项

标准编制组开展了大量前期资料收集与研究。结合专家共识和指南要求，讨论并确定了标准内容和框架结构，完成项目建议书，并于2024年4月正式立项。

### （二）成立标准编制组

2024年4月，深圳市卫生健康发展研究与数据管理中心牵头正式成立了标准编制组，并组织各参与编制人员召开第一次标准讨论会，在结合国内外相关指南、共识、文献基础上，进一步明确了生成式人工智能应用指南框架要点内容。

### （三）文献研究

编制组系统广泛搜集生成式人工智能相关政策、法律法规、国内外相关标准，科研文献。

#### （四）调查研究

2024 年 4 月至 5 月，标准编制组开展了前期资料收集与研究工作。编制组实地走访、访谈调研福田区妇幼保健院，深圳市人民医院，蛇口区人民医院等 6 家医院，了解生成式人工智能应用情况及相关问题。并对 58 家市区级医院生成式人工智能应用情况进行问卷调查，其中医学影像辅助诊断应用 16 家，智能导诊应用 14 家，临床辅助决策支持应用 10 家。

#### （五）标准草案编制

2024 年 6 月—7 月，编制组根据前期文献、实地调研，专家访谈情况，编制标准草案。

#### （六）标准研讨修改

2024 年 8 月，标准编制组召开专题讨论会，邀请标准化、医疗人工智能专家、医疗管理人员、医疗信息专家，进行标准框架及内容研讨，根据专家意见，编制组修改完善了标准基本框架，并对标准内容进行了论证，确定标准的关键评价原则和框架。

2024 年 11 月，标准编制组开展标准主要技术内容研讨会，医院管理、临床医务人员、人工智能专家、医疗信息化人员等专家领域，进行标准修改，明确下一步工作进度安排。

#### （七）征求意见

深圳市卫生健康委、深圳市卫生健康发展研究和数据管理中心于 2025 年 3 月、6 月，征集各区卫生健康行政主管部门、市属医疗卫生机构等相关单位意见，收到 48 家单位回

复，共计 45 条意见，其中 39 家单位无意见。采纳 30 条，部分采纳 7 条，不采纳 8 条，对部分采纳意见和不采纳意见陈述了理由。项目组根据意见进一步修改和完善本文件送审稿，并提交至深圳市卫生健康委员会。

### 三、确定标准主要内容的依据以及与国内领先、国际先进标准的对标情况

（一）本标准的制订根据《中华人民共和国标准化法》和《深圳经济特区健康条例》及有关法规、规章的原则要求进行编写。

（二）与本标准相关的法律、法规、标准：主要有 2023 年《生成式人工智能服务管理暂行办法》，2021 年世界卫生组织发布的《Ethics and governance of artificial intelligence for health》，《人工智能 大模型 第 1 部分：通用要求》（GB/T 45288.1-2025），《人工智能 大模型 第 2 部分：评测指标与方法》（GB/T 45288.2-2025），《人工智能 大模型 第 3 部分：服务能力成熟度评估》（GB/T 45288.3-2025）。标准中已有相关规定的，本标准予以吸收和引用。

### 四、主要条款的说明

#### 1 范围

本文件提供了卫生健康领域生成式人工智能应用的原则、启动、评估、分类、部署、运维等指南。本文件适用于指导深圳市各机构在卫生健康领域应用生成式人工智能的活动。

## 2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；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 35273 信息安全技术 个人信息安全规范

GB/T 39725 信息安全技术 健康医疗数据安全指南

GB/T 39786 信息安全技术 信息系统密码应用基本要求

## 3 术语和定义

根据对相关标准、文献的研究，本标准给出了生成式人工智能、使用者、服务提供者相关的术语及其定义。

## 4 原则

生成式人工智能原则包括合规性、安全性、透明性、公平性、先进性、可控性。

## 5 启动

成立管理小组，管理小组由具有不同专业及背景的人员组成；管理小组进行需求分析和目的确定，确定生成式人工智能初步模型选择。

## 6 评估

为了有效地评估生成式人工智能的实施情况，需要考虑或测试性能、安全、伦理、适配度、成本收益、用户体验六个维度。

## 7 分类

根据模型管理和使用需求，遵循关键要素的评估结果，从场景应用角度进行分类，将其分为不应用，条件性应用，一般性应用三个类别。

## **8 部署**

根据模型评估与分类原则，制定管理制度，对模型进行调整，在部署前进行评估验证，部署后进行集成测试及安全测试，确保模型性能及安全运行。

## **9 运维**

根据建立的管理机制，持续监测模型性能。定期更新模型。定期进行核查、评估，并收集反馈，持续改进和完善模型的功能。终端用户建立明确的服务和使用条款，并定期进行员工培训。对算法、数据、伦理、行为进行审计追溯。

## **五、是否涉及专利等知识产权问题**

无。

## **六、重大意见分歧的处理依据和结果**

无。

## **七、实施标准的措施建议**

标准正式发布后会同标准化技术机构，卫生健康领域应用生成式人工智能机构及相应业务主管部门，通过培训班、讲座、微信文章、新闻报道等多种渠道、多种方式进行宣贯，使相关机构尽快熟悉、了解、掌握和应用本标准；并根据应用情况，进一步修订完善标准、深化应用。

## **八、其他需要说明的事项**

无。