

# 2023医疗健康AI大模型行业研究报告

亿欧智库 <https://www.iyiou.com/research>

Copyright reserved to EO Intelligence, December 2023

## 前言

- 2023年9月，习近平总书记在黑龙江考察调研期间首次提到“新质生产力”，强调整合科技创新资源，引领发展战略性新兴产业和未来产业，加快形成新质生产力。
- 医疗行业是一个典型的知识和技术密集型行业，其发展水平直接关系到国民健康和生命质量。医疗健康AI大模型，作为人工智能的一个分支，能够通过学习大量的数据来生成新的数据实例，这在医疗领域有着广泛的应用前景，如药物研发、医学影像、医疗文本分析等。这些应用不仅能够提升医疗服务的质量和可及性，还能够推动整个医疗产业的创新发展，形成新的产业生态和生产力。因此，“新质生产力”概念下的科技创新资源整合，对于医疗和生成式AI领域的发展至关重要。
- 基于此，亿欧撰写了本报告，旨在分析当前医疗健康AI大模型在实际应用中的成效，提高医疗各界对医疗健康AI大模型的认识和理解，讨论其对医疗服务可及性和质量的潜在影响，增强医疗各界对新技术的接受度。

## 目录

## CONTENTS

### 01 医疗健康AI大模型发展背景

- 医疗健康AI大模型发展过程
- 医疗健康AI大模型的技术支持
- 医疗健康AI大模型在医疗场景下的初步表现

### 02 医疗健康AI大模型应用场景

- 医疗健康AI大模型应用场景盘点
- 医疗健康AI大模型医疗健康场景应用-医学影像
- 医疗健康AI大模型医疗健康场景应用-药物研发
- 医疗健康AI大模型健康场景应用-医疗文本处理
- 医疗健康AI大模型健康场景应用-学术科研

### 03 国内医疗健康AI大模型市场发展简析

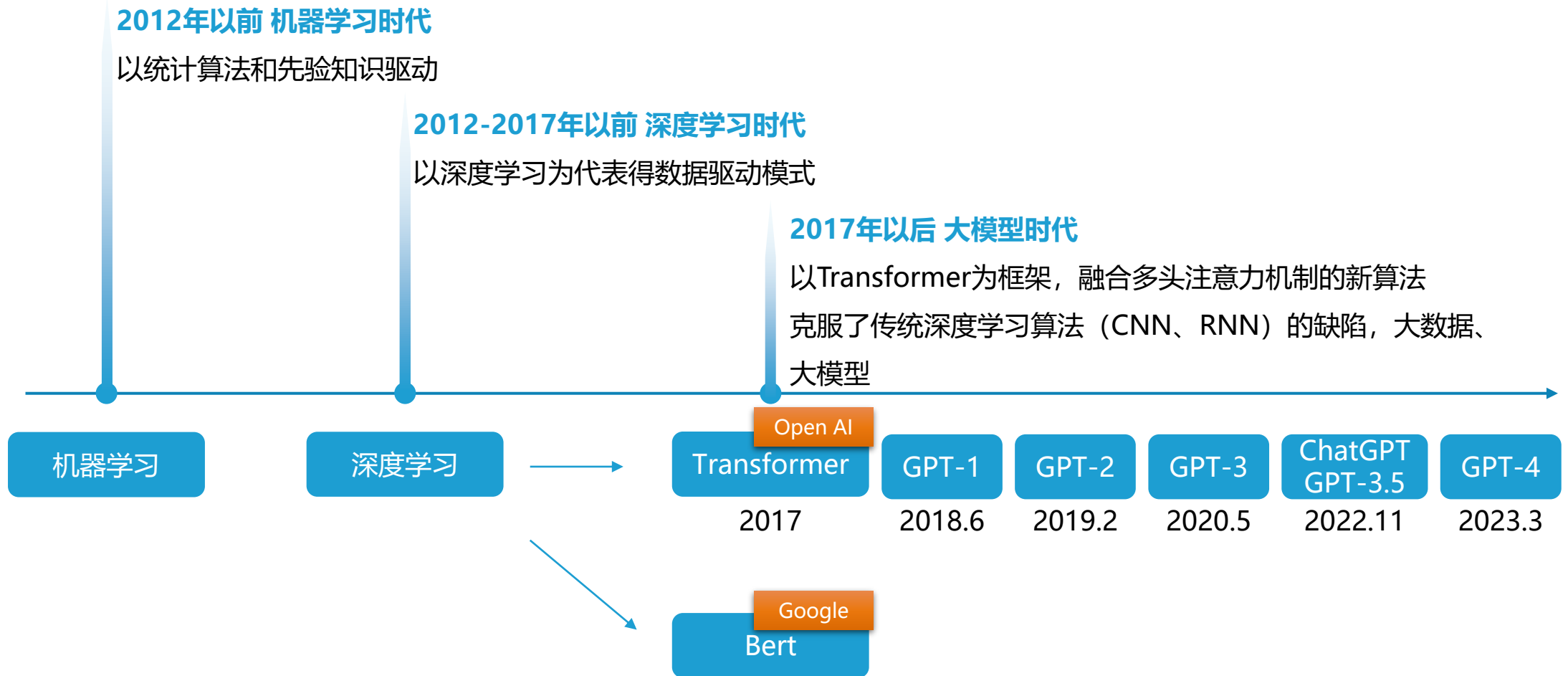
- 医疗健康AI大模型产品市场规模及增速
- 医疗健康AI大模型产品应用落地的企业类型
- 医疗健康AI大模型产品及企业图谱医疗
- 医疗健康AI大模型合作梳理

### 04 医疗健康AI大模型平台测试结果分析

- 医疗健康AI大模型测试背景
- 医疗健康AI大模型测试核心发现
- 医疗健康AI大模型测试结果总结

# 医疗健康AI大模型技术发展过程：从深度学习过渡到大模型时代

- ◆ 在1956年的达特茅斯会议上，“人工智能”的概念被首次提出，六十多年以来，历经逻辑推理、专家系统、深度学习等技术的发展，人工智能也经历了数次沉浮，有如日中天的发展时期，也有因技术不足热度退去后的寒潮。而每一次寒潮后，也会经历技术的蛰伏发展，为人工智能带来里程碑式的进步。



# 众多高校研发医疗健康AI大模型，为行业发展的技术底座提供了坚实的基础

- ◆ 目前众多高校启动了关于中文医疗健康AI大模型的研发并在Github上发布源代码，增加了商业企业对于医疗健康AI产品研发的技术支持。
- ◆ 从研发模型的功能分析，大多围绕问诊对话、病历结构化等文字处理相关的功能。

## 亿欧智库：开源中文医疗健康AI大模型研发信息

研发高校信息	模型名称	模型功能
哈尔滨工业大学社会计算与信息检索研究中心健康智能组	本草BenTaso	通过医学知识图谱和GPT3.5 API构建了中文医学指令数据集，并在此基础上对LLaMA进行了指令微调，提高了LLaMA在医疗领域的问答效果。此外，项目还尝试利用GPT3.5 API将医学文献中的【结论】作为外部信息融入多轮对话中，在此基础上对LLaMA进行了指令微调。
华东师范大学计算机科学与技术学院智能知识管理与服务团队	ChatMed	医疗问诊
	ShenNong-TCM-LLM - 神农中医药大模型	医疗问诊
澳门理工大学	CareLlama关怀羊驼中文医疗大模型	通用大模型
	MedQA-ChatGLM	医学对话
	XrayGLM	首个会看胸部X光片的中文多模态医学大模型
华南理工大学未来技术学院广东省数字孪生人重点实验室合作单位包括广东省妇幼保健院、广州市妇女儿童医疗中心和中山大学附属第三医院等	BianQue扁鹊 - 中文医疗对话大模型	中文医疗对话
	SoulChat灵心健康大模型	中文医疗对话
华东理工大学信息科学与工程学院	孙思邈中文医疗大模型	医疗问诊
	Mindchat漫谈中文心理大模型	中文医疗对话-心理咨询、心理评估、心理诊断、心理治疗
浙江大学、网新数字健康联合研究中心	QiZhenGPT启真医学大模型	药品知识问答，医患问答、病历自动生成
上海人工智能实验室、上海交通大学-清源研究院、华东理工大学-自然语言处理与大数据挖掘实验室	PULSE中文医疗大模型	健康教育、医师考试问题、报告解读、医疗记录结构化以及模拟诊断和治疗
上海交通大学未来媒体网络协同创新中心和上海人工智能实验室智慧医疗中心	MING	基于共计28科室的中文医疗共识与临床指南文本，从而生成医疗知识覆盖面更全，回答内容更加精准的高质量指令数据集。
上海交通大学	DoctorGLM	中文问诊模型
中国科学院自动化研究所	紫东太初2.0	智能化疾病管理、医疗多模态鉴别诊断
上海人工智能实验室	Open-MEDLab	医疗多模态基础模型群

- ◆ 数据库的建设对健康医疗大数据的落地应用的至关重要，在医疗领域建立权威、科学、规范的临床重点专科标准数据库，能够支持AI大模型医疗健康场景应用落地以及解决健康需求面临的挑战。
- ◆ 2016年，国家卫健委能力建设和继续教育中心开始着手建立数据库，由于中心拥有非常优质的专家资源，保证了数据质量。在数据库建立的过程中，也遇到了非常多的无法解决的问题，而我国将临床专家的智慧发挥到了极致，即根据临床需求来反推中心需要进行的工作。由此，中心制定了入库收集的标准，收集数据，并将其应用起来。由于在数据库建设中，中心一直坚持以临床需求应用为导向，使建立过程虽然历经艰辛但依然坚持走到了现在。

## 亿欧智库：人工智能应用公开数据集



# 医疗健康AI大模型在医疗场景下的初步表现：在问答的准确性和全面性有优势，但资料搜集方面可用性较低

- ◆ 下表的数据反映了不同研究中对ChatGPT在不同场景下的表现和结果进行评估的情况。可以看出，不同研究对ChatGPT的评价结果存在差异。
- ◆ 在回答患有前列腺癌患者的问题时，ChatGPT的回答准确性和全面性较高，易读性和稳定性也较好；在回答有关肝硬化和肝细胞癌的问题时，ChatGPT的准确率和全面性相对较高；**在检索临床放射学的信息和参考文献时，ChatGPT的回答准确率为67%，但参考文献的可用性较低**；在在线医生咨询中，ChatGPT的回答长度较长，质量较高，且被评为富有同理心。

## 亿欧智库：ChatGPT在不同医疗场景下的反馈结果

年份	场景	结果
2023年	回答患有前列腺癌患者的问题	ChatGPT Plus版：准确性：“正确” = 93.94%；全面性：“非常全面” = 91.94%；易读性：“非常易懂” = 100%；稳定性：一致性 = 90.91%
2023年	回答有关肝硬化和肝细胞癌的问题	正确率：肝硬化（79.1%），肝细胞癌（74.0%） 全面性：肝硬化（47.3%），肝细胞癌（41.1%）
2023年	检索临床放射学的信息和参考文献	回答：67%正确，33%有错误 参考文献：36.2%可在互联网上找到，63.8%由ChatGPT生成
2023年	美国医学执照考试	ChatGPT在4个数据集（AMBOSS-Step1，AMBOSS-Step2，NBME-Free-Step1和NBME-Free-Step2）中的准确率分别为44%（44/100），42%（42/100），64.4%（56/87）和57.8%（59/102）。
2023年	在线医生咨询中的公共对话	回答内容长度：平均（IQR）：医生：52 [17-62]个单词；ChatGPT：211 [168-245]个单词；t= 25.4；P < 0.001。 回答质量：比例：医生：22.1%，95%置信区间：16.4%-28.2%；ChatGPT：78.5%，95%置信区间：72.3%-84.1%。回答被评为富有同理心：

## 目录

## CONTENTS

### 01 医疗健康AI大模型发展背景

- 医疗健康AI大模型发展过程
- 医疗健康AI大模型的技术支持
- 医疗健康AI大模型在医疗场景下的初步表现

### 02 医疗健康AI大模型应用场景

- 医疗健康AI大模型应用场景盘点
- 医疗健康AI大模型医疗健康场景应用-医学影像
- 医疗健康AI大模型医疗健康场景应用-药物研发
- 医疗健康AI大模型健康场景应用-医疗文本处理
- 医疗健康AI大模型健康场景应用-学术科研

### 03 国内医疗健康AI大模型市场发展简析

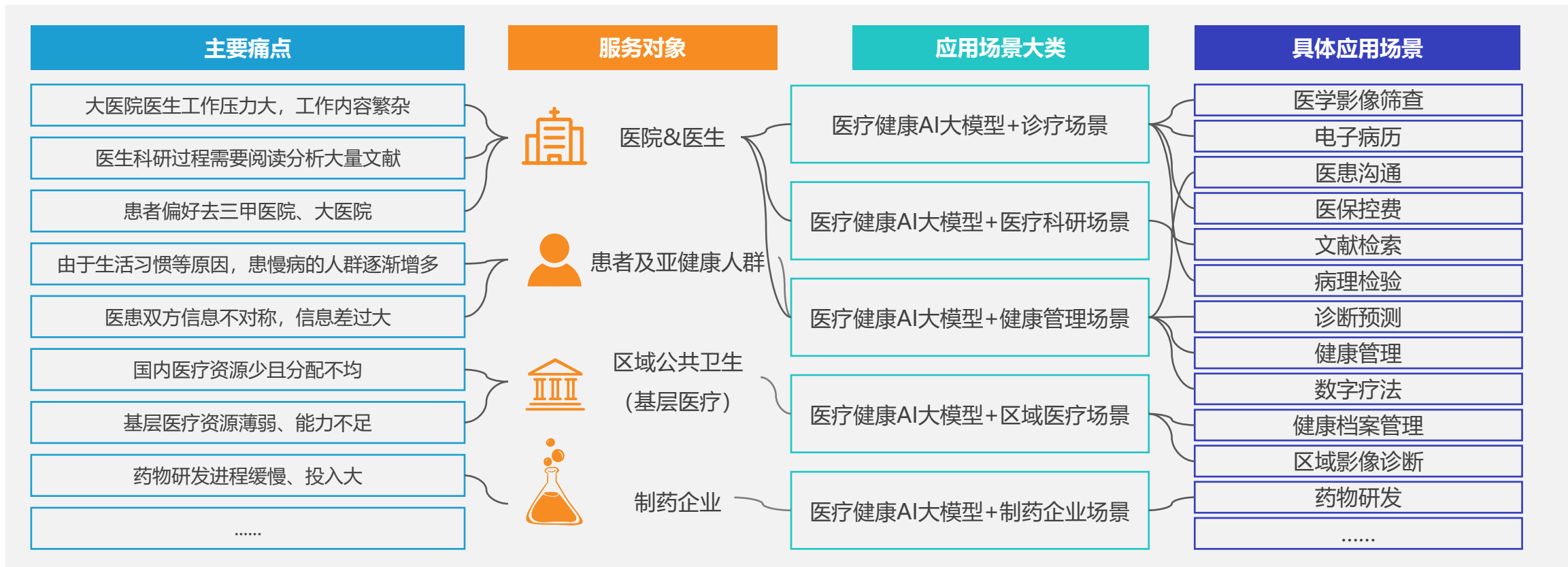
- 医疗健康AI大模型产品市场规模及增速
- 医疗健康AI大模型产品应用落地的企业类型
- 医疗健康AI大模型产品及企业图谱医疗
- 医疗健康AI大模型合作梳理

### 04 医疗健康AI大模型平台测试结果分析

- 医疗健康AI大模型测试背景
- 医疗健康AI大模型测试核心发现
- 医疗健康AI大模型测试结果总结



## 亿欧智库：医疗健康AI大模型应用场景



阿里巴巴  
通义仁心

腾讯  
混元

京东  
京医千询

百度  
文心

华为  
盘古

商汤大医

智谱AI

百川智能

- ◆ 医学影像和大模型结合的优势在于提升了疾病诊断的准确性和效率，为医生提供了更全面的信息和更好的治疗方案。但医学影像和大模型结合也存在局限性例如，数据获取困难、兼容性问题、复杂性和可解释性不足。

## 亿欧智库：医疗健康AI大模型在医学影像中的作用

### 差异诊断

帮助放射科医生对发现的病变进行差异诊断，并提供额外的影像技术信息来确诊。例如，ChatGPT可以提供关于潜在的骨肿瘤的信息，同时也可以为罕见病变的患者提供有效的随访建议。

### 临床决策支持

为放射科医生提供临床决策支持。通过提供相关的医学知识和建议，ChatGPT可以帮助医生制定治疗方案、选择适当的影像技术，并提供随访建议等。

### 教育和培训

对于经验不足或非专科的放射科医生来说，医疗健康AI大模型可以作为一个有益的教育和培训工具。它可以提供有关疾病诊断和处理的信息，帮助医生学习和提高技能。

### 医学影像中应用案例

问ChatGPT：一个十岁的女孩进行了X射线检查，发现股骨近干骺端皮质旁边有硬化区域。患者的大腿有疼痛。你的诊断是什么？

ChatGPT提供了关于潜在骨肿瘤的信息，并表示可能需要进行CT或MRI来确认诊断。因此，**ChatGPT可以帮助放射科医生决定是否需要进行CT扫描以获得更准确的诊断或评估。这种联系使得ChatGPT成为一个有用的工具，可以在放射科医生的临床实践中提供有关影像技术选择的建议。**

### 代表案例：深睿医疗通用大模型全面开启全场景全流程AI数智化新时代

深睿医疗智慧影像&大数据通用平台-Deepwise MetAI，融合计算机视觉、NLP、深度学习、大模型等人工智能前沿技术，将影像科日常应用产生的数据结构化、形成优质的数据资产，可以在技师、医生、科室的管理者之间自由流通、实现重建、打印、诊断、会诊、教学、科研的一站式影像数智化，更好地支持诊疗流程、科学研究、科室建设乃至医院整体业务的运转，打造新一代人工智能数据应用新基座，全面助力医院智慧化高质量发展。MetAI，遇见AI，一切皆有可能！

### 代表案例：联影智能多模态大模型涌现式跃迁与突破



医疗影像大模型：基于多病种影像数据，训练医疗影像基础模型，可快速衍生到其他新病种



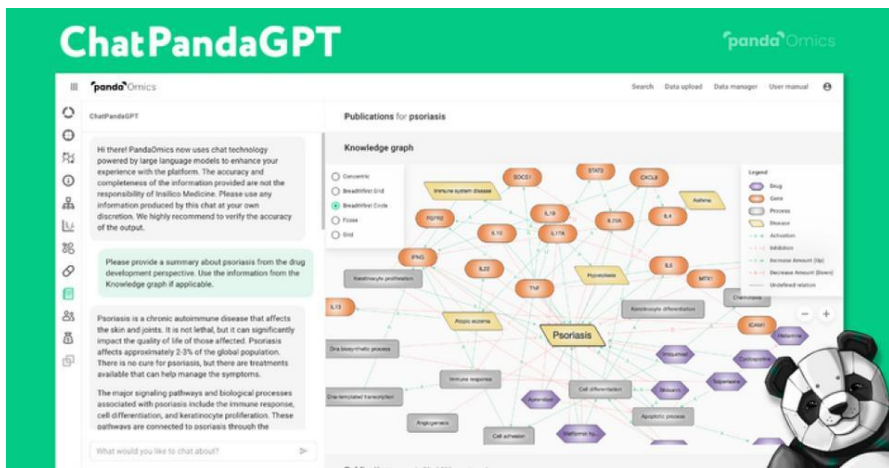
医疗文本大模型：用医疗文本对通用语言大模型进行调优，构建医疗文本/知识模型



医疗混合模态大模型：构建医疗场景中融合多模态信息（影像、文本、视频）的大模型

◆ 大模型和药物研发的结合可以提高药物研发效率、降低成本、提高准确性和个性化治疗，是药物研发领域的重要趋势和发展方向。

## 亿欧智库：Insilico Medicine发布ChatPandaGPT和知识图谱



- Insilico Medicine公司将最新版本的高级AI聊天工具ChatGPT集成到其PandaOmics平台中，从而实现了名为"ChatPandaGPT"的新功能。
- 该功能使得研究人员能够**更加高效、准确、个性化地访问有关分子生物学、治疗靶点发现和药物开发的信息。**
- Insilico Medicine还发布了一个**基于Transformer的知识图谱，从科学出版物列表中提取相关信息，并生成一个视觉上吸引人的图表，将基因、疾病、化合物和生物过程联系起来，提供了对疾病分子机制以及疾病背景下基因和化合物生物活性的见解。**

## 亿欧智库：大模型vs传统深度学习在生成人工蛋白质的效果对比

**ProGen模型是一个基于Transformer架构的大模型，拥有12亿个可训练参数。**在一项关于ProGen的模型中，研究者发现：与通用的深度学习算法进行对照试验后，ProGen这种专门设计的语言模型可以生成具有可预测功能的蛋白质序列，并且在不同蛋白质家族上具有广泛的适用性。

	ProGen模型	传统深度学习
数据量需求	<ul style="list-style-type: none"> <li>• ProGen是一种基于条件语言模型的蛋白质生成模型，它通过学习大规模的蛋白质序列数据来生成新的蛋白质序列。</li> <li>• <b>ProGen的数据需求量相对较小，使用了280万个蛋白质序列进行训练，</b></li> </ul>	<ul style="list-style-type: none"> <li>• 深度学习模型具有大量的参数，需要足够的约束和优化这些参数。因此，复杂任务深度学习模型需要非常多的样本来达到良好性能。</li> <li>• <b>因为参数众多，需要足够数据来优化模型。所需的数百万到数十亿个样本。</b></li> </ul>
控制能力	<ul style="list-style-type: none"> <li>• ProGen是一种条件语言模型，可以根据输入的控制标签生成具有特定功能的蛋白质序列。</li> <li>• ProGen模型可以更好地捕捉到这些家族的序列特征和分布，从而提高生成性能。</li> </ul>	<ul style="list-style-type: none"> <li>• 深度学习模型是一种通用模型，在生成特定家族或功能的蛋白质序列方面可能存在一定的限制。</li> <li>• 深度学习在蛋白质生成和设计任务中可能<b>难以学习到特定蛋白质家族的特征和分布。</b></li> </ul>
生物物理建模	<ul style="list-style-type: none"> <li>• ProGen模型的生成结果可以与生物物理建模相结合，这种结合可以<b>提供对生成蛋白质的更深入理解和评估。</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>深度学习通常难以直接预测蛋白质的结构和功能，生成的蛋白质的结构和功能验证仍然需要结合生物物理建模等方法进行进一步的验证。</b></li> </ul>

◆ 使用大型语言模型处理医疗文本可以为医疗专业人员提供更准确、快速、个性化的诊断和治疗建议，从而提高医疗效果和患者满意度。

## 亿欧智库：医疗健康AI大模型应用于医疗文本处理的具体场景

### 1 临床记录自动化

将医生的口述或非结构化的记录转换为结构化的电子健康记录

### 2 医疗文献分析

从大量医学文献中提取关键信息，帮助研究人员快速了解研究趋势和发现。

### 3 病例报告和研究摘要的生成

自动生成病例报告和研究摘要，减少医疗专业人员的文档负担。

### 4 自动问答系统

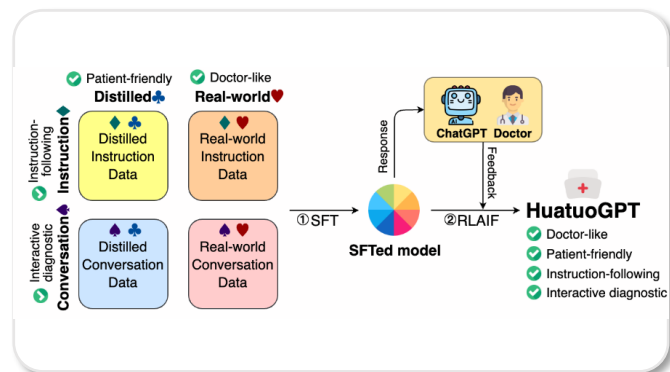
为患者和医疗专业人员提供基于证据的医疗信息，支持临床决策。

### 5 医疗编码和计费

帮助将临床服务转化为标准化的医疗计费代码。



## 应用案例



## 开源中文医疗大模型华佗GPT

香港中文大学（深圳）和深圳市大数据研究院的王本友教授团队开发了HuatuoGPT医疗大模型。它结合了ChatGPT和医生回复的数据，**使语言模型具备像医生一样的诊断能力和提供有用信息的能力，同时保持对用户的流畅交互和内容丰富性，让对话更加顺畅。**

## 亿欧智库：医疗AI大模型应用于学术科研的具体场景

- 01 生成初稿**  
ChatGPT和LLMs可以用于生成学术论文、研究提案或其他学术文件的初稿。它们可以为作者提供一个起点，帮助加快写作过程。
- 02 编辑和完善文本**  
ChatGPT和LLMs可以用于编辑和完善学术文本。它们可以提出风格、语法和清晰度方面的改进意见，帮助作者提高写作质量。
- 03 文献综述和数据分析**  
LLM可以帮助研究人员找到相关研究、提取关键信息和分析大型数据集，节省时间和精力。
- 04 语言翻译**  
LLMs可以用于语言翻译，使研究人员能够访问和理解不同语言的科学文献。这有助于促进国际间的合作和知识交流。
- 05 临床决策支持**  
LLMs可以在临床决策中提供支持，通过分析患者数据、医疗记录和研究文献，帮助医疗专业人员诊断疾病。
- 06 医学教育和培训**  
LLMs可以在医学教育和培训项目中使用。它们可以提供互动学习体验，回答学生的问题，并模拟患者接触，增强学习过程。

## 亿欧智库：医疗AI大模型应用于学术科研的风险与挑战

- 01 可信度挑战**  
由于ChatGPT是基于大量文本数据进行训练的语言模型，它生成的内容可能缺乏可信度。它可能会产生不准确、不完整或有偏见的信息，这可能会对医学写作的准确性和可靠性产生负面影响。
- 02 抄袭风险**  
在医学写作中，抄袭是严重违反学术道德和版权法的行为。使用ChatGPT时，作者需要确保生成的内容不会侵犯他人的版权，并采取适当的引用和参考文献的方式。
- 03 透明度挑战**  
ChatGPT生成的文本可能缺乏透明度，即无法追溯其来源和生成过程。这可能导致读者无法确定文本的可靠性和准确性。在医学写作中，读者需要了解文本的来源和基础数据，以便评估其可信度。
- 04 责任和权威性挑战**  
在医学写作中，作者需要对他们使用ChatGPT生成的内容负责，并确保其准确性和权威性。医学写作需要经过专业人士的审查和验证，以确保内容的准确性和可靠性。
- 05 隐私和数据保护风险**  
使用ChatGPT时，需要确保患者和个人的隐私和数据保护。生成的文本可能包含敏感信息，因此需要采取适当的措施来保护数据的安全性和隐私。

- ◆ 2023年6月27日，北京市科委、中关村管委会举办首期北京市人工智能策源地引领推介活动，发布了北京市首批人工智能行业大模型应用案例。
- ◆ 北京是国内大模型创新基础最好、人才团队最集中、研发能力最强、产品迭代最活跃的地区。一大批一线科学家和创业者集聚北京，涌现出文心一言、ChatGLM等一批具有代表性的大模型产品。科技部新一代人工智能发展研究中心发布的《中国人工智能大模型地图研究报告》显示，我国10亿参数规模以上的大模型已发布79个，其中北京发布的数量最多，共38个，占全国近一半，大模型学者指数、模型开源数量和影响力等指标均位居全国第一。首批应用案例中，涵盖两个医疗相关模型。

## 北京智谱华章科技有限公司和北京中医药大学东方医院共同开发的数字中医大模型示范应用



在智慧医疗领域，基于智谱华章高精度千亿中英双语稠密模型“GLM-130B”，面向中医领域名医经验挖掘整理需求，构建数字中医服务平台，探索高危肺结节人工智能临床诊疗和临床评价研究等解决方案，实现中医临床经验的智慧化复制新模式。

## 云知声科技股份有限公司和北京友谊医院共同开发的基于山海大模型的门诊病历生成系统示范应用



在智慧医疗领域，基于云知声700亿参数规模的自研“山海”大模型，结合前端声音信号处理、声纹识别、语音识别、语音合成等全栈式智能语音交互技术，预计可提升医生的电子病历录入效率超过400%，节约单个患者问诊时间超过40%，提升医生门诊效率超过66%。

## 目录

## CONTENTS

### 01 医疗健康AI大模型发展背景

- 医疗健康AI大模型发展过程
- 医疗健康AI大模型的技术支持
- 医疗健康AI大模型在医疗场景下的初步表现

### 02 医疗健康AI大模型应用场景

- 医疗健康AI大模型应用场景盘点
- 医疗健康AI大模型医疗健康场景应用-医学影像
- 医疗健康AI大模型医疗健康场景应用-药物研发
- 医疗健康AI大模型健康场景应用-医疗文本处理
- 医疗健康AI大模型健康场景应用-学术科研

### 03 国内医疗健康AI大模型市场发展简析

- 医疗健康AI大模型产品市场规模及增速
- 医疗健康AI大模型产品应用落地的企业类型
- 医疗健康AI大模型产品及企业图谱医疗
- 医疗健康AI大模型合作梳理

### 04 医疗健康AI大模型平台测试结果分析

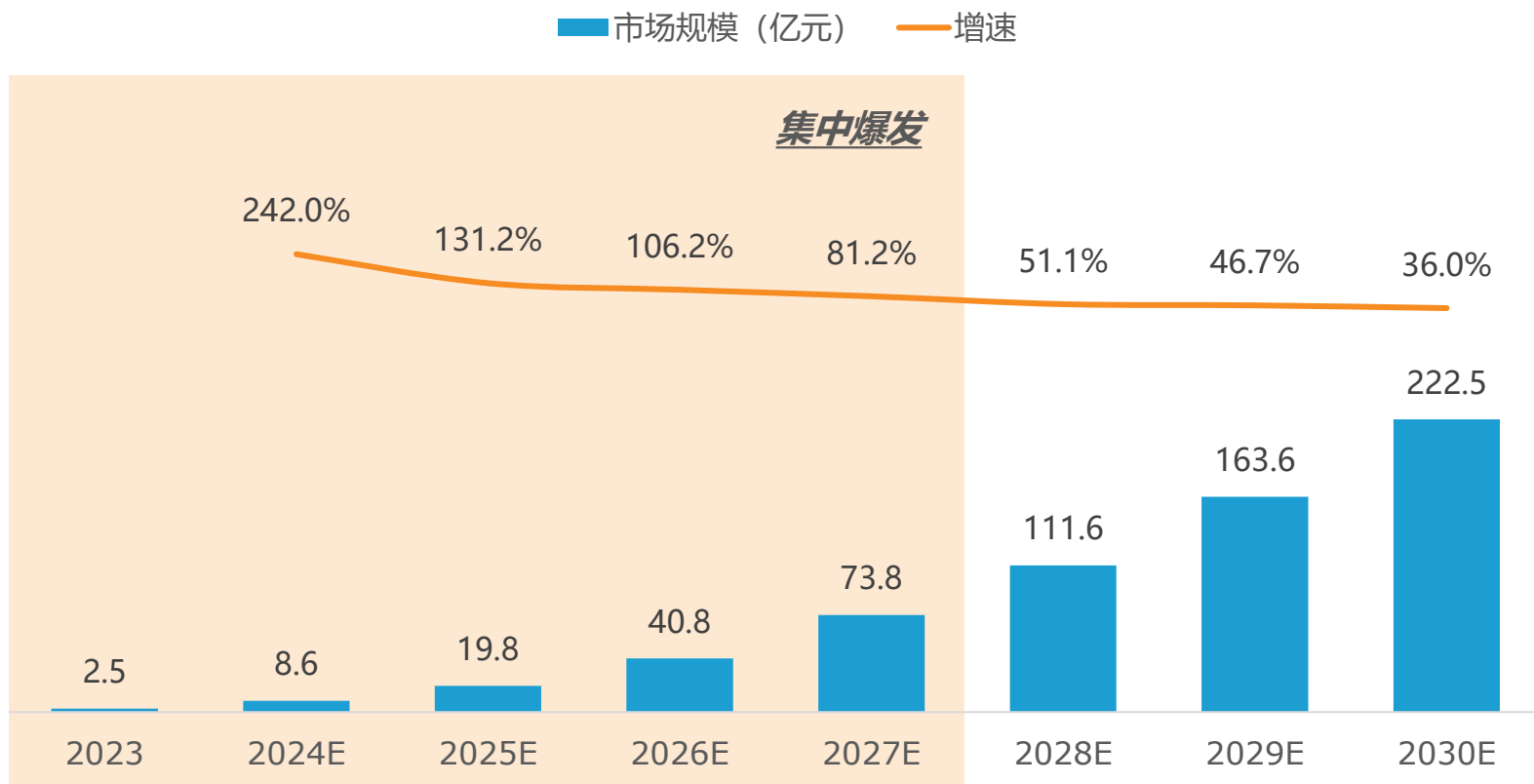
- 医疗健康AI大模型测试背景
- 医疗健康AI大模型测试核心发现
- 医疗健康AI大模型测试结果总结

# 原有医疗场景应用创新的难度被大大降低，预计2023-2027年为医疗健康AI大模型集中爆发阶段

◆ 随着技术成熟度以及数据可用性增加，医疗健康AI大模型对医疗场景中（如药物发现、个性化医疗、医学影像、数据增强）的部分研发阻碍给出解决方案，加快医疗研发速度。同时，其生成式人工智能操作难度低，更能够被医疗领域的专业人员接受并使用。因此根据亿欧预测，2023-2027年为医疗健康AI大模型集中爆发的阶段。

◆ 虽然医疗健康AI大模型在一定程度上降低了某些方面的成本，但它们的使用和维护仍然需要相当的资源投入，特别是对于需要高度定制化和专业化的应用场景。

### 亿欧智库：2023-2030年中国医疗健康AI大模型产品市场规模及增速



爆发原因一：

**研发效率大大提升。**医疗健康AI大模型能够自动生成内容、代码、数据分析报告等。这可以帮助开发者和医疗科学家快速原型开发和测试，从而减少时间和成本的投入。

爆发原因二：

**更易于接入的API服务。**很多公司提供了基于这些医疗健康AI大模型的API服务，这意味着即使是没有深厚技术背景的用户也可以通过简单的API调用来利用AI的能力。

爆发原因三：

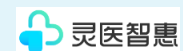
**大型预训练模型可以被开发者直接使用。**这减少了从头开始训练模型的需求。开发者可以通过微调这些模型来适应特定的应用场景，这大大降低了数据收集和计算资源的成本。



## 亿欧智库：医疗健康AI大模型产品应用落地的企业类型



## 患者问诊全流程



百度·  
灵医大模型



腾讯·  
混元大模型



阿里·  
通义仁心



京医千询



星火大模型



商汤大医



左手医生·  
左医GPT



云知声·  
山海大模型



微脉·  
CareGPT



春雨慧问



轻松健康  
轻松问医



HealthGPT

## 医生助手/医院管理



东软·  
添翼大模型



YiduCore  
大模型



WiNEX Copilot



MEDGPT



白泽晓医疗  
大模型



百度·  
灵医大模型



BSoftGPT  
大模型



ClouDGPT  
大模型



紫薇平台  
AI助手

## 中医



智谱华章·  
数字中医  
大模型



岐黄问道  
大模型

## 医学影像



Deepwise MetAI

## 药物研发



盘古大模型

## 妇科健康科普



爱巢GPT

# 医疗科技企业与基础大模型企业积极合作，AI大模型医疗健康生态正在逐步建立

◆ 当前医疗科技企业与基础大模型企业合作数量较多，双方借助各自在科技及医疗的优势，打造行业生态。

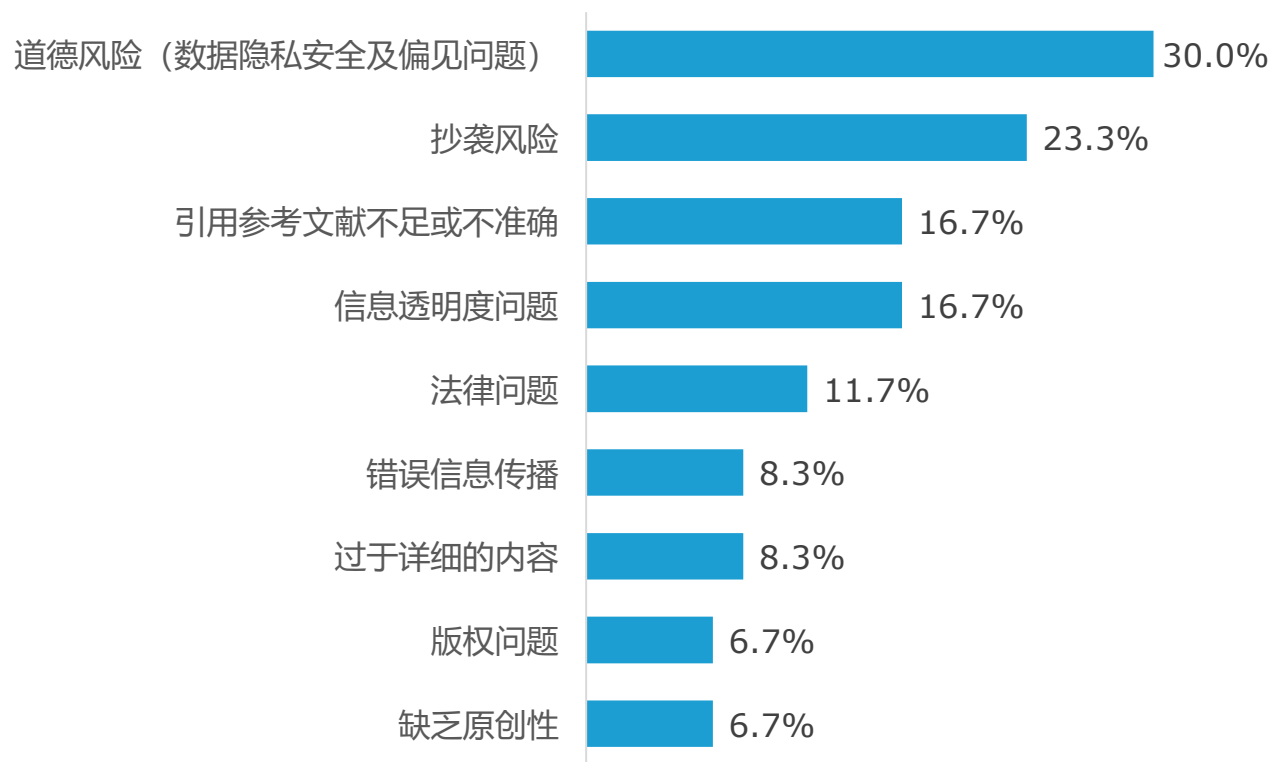
## 亿欧智库：2023年1-11月医疗健康AI大模型部分企业或组织间合作梳理

合作方式	发布时间	合作内容
医疗科技企业+基础大模型企业	2023年7月9日	<b>医渡科技&amp;昇腾AI</b> ：医渡科技携手华为启动大模型联合创新
	2023年6月7日	<b>润达医疗&amp;华为云</b> ：润达医疗与华为云战略合作 打造普惠AI医疗服务大模型
	2023年5月8日	<b>嘉会医疗&amp;商汤科技</b> ：嘉会医疗携手商汤科技，AI赋能打造国际化智慧医院样板
	2023年7月10日	<b>金域医学&amp;华为云</b> ：金域医学携手华为云打造医检行业大模型
	2023年2月19日	<b>文心一言&amp;有来医生/莫廷医疗/启灏医疗/医学界/东软医疗</b> ：百度牵手多家医疗企业，类ChatGPT模式加速医疗产业升级
	2023年7月10日	<b>万达信息&amp;华为云</b> ：万达信息与华为云共同签署盘古大模型合作协议
	2023年7月20日	<b>灵医智惠&amp;固生堂</b> ：灵医智惠与固生堂联合开展大模型战略合作
	2023年9月8日	<b>腾讯健康&amp;迈瑞医疗</b> ：迈瑞医疗与腾讯健康达成战略合作
	2023年11月1日	<b>腾讯健康与圆心科技</b> ：腾讯健康与圆心科技举行战略合作签约
医疗科技企业+医院	2023年7月7日	<b>联影智能&amp;中山医院</b> ：联影智能携手中山医院，打造全病程智医诊疗大模型
科研机构、高校及医院合作	2023年6月29日	全球首个医疗多模态基础模型群 “OpenMEDLab浦医”

# 医疗健康AI大模型的应用仍存在诸多风险，将有可能阻碍其商业化路径，国内外陆续推出相关政策

- ◆ 医疗大模型的商业化面临数据隐私、知识产权和道德责任等挑战，会严重损害企业声誉并带来经济损失，增加了运营成本，这可能会阻碍大模型的商业化。
- ◆ 为了克服这些障碍，企业需要采取措施加强数据保护、获得必要的版权授权、确保模型的公正性和合规性，同时通过提高透明度和公众教育来建立用户信任。这要求在技术创新和社会责任之间找到一个平衡点，确保医疗大模型能够以负责任和可持续的方式推向市场。

## 亿欧智库：使用医疗大模型在医疗场景中存在的风险



### 国内外推出相关政策，希望推动医疗大模型平稳落地

- 2023年12月8日，欧盟就全球首个全面监管AI的《人工智能法案》达成初步协议，该协议对所有通用人工智能模型都提出了透明度要求，对更强大的模型则提出了更严格的规定。
- 在我国推出的《国家限制类技术临床应用管理规范（2022年版）》针对使用人工智能辅助治疗技术，在人员配备、培训等多方便做出了规定。

### 构建安全可信的人工智能：数据保护、透明度与伦理公平性

- 加强数据加密和匿名处理，确保符合数据保护法规；建立透明度和问责制度，增强用户信任；在模型开发和部署过程中考虑到伦理和公平性，确保模型的公正性。

## 目录

## CONTENTS

### 01 医疗健康AI大模型发展背景

- 医疗健康AI大模型发展过程
- 医疗健康AI大模型的技术支持
- 医疗健康AI大模型在医疗场景下的初步表现

### 02 医疗健康AI大模型应用场景

- 医疗健康AI大模型应用场景盘点
- 医疗健康AI大模型医疗健康场景应用-医学影像
- 医疗健康AI大模型医疗健康场景应用-药物研发
- 医疗健康AI大模型健康场景应用-医疗文本处理
- 医疗健康AI大模型健康场景应用-学术科研

### 03 国内医疗健康AI大模型市场发展简析

- 医疗健康AI大模型产品市场规模及增速
- 医疗健康AI大模型产品应用落地的企业类型
- 医疗健康AI大模型产品及企业图谱医疗
- 医疗健康AI大模型合作梳理

### 04 医疗健康AI大模型平台测试结果分析

- 医疗健康AI大模型测试背景
- 医疗健康AI大模型测试核心发现
- 医疗健康AI大模型测试结果总结

# 医疗健康AI大模型测试背景

自大模型受到广泛关注以来，出现了较多关于医疗大模型的测试文章，但测试大多围绕GPT在回答美国执业医师资格考试（USMLE）的结果进行。

在此背景下，亿欧围绕国内医疗大模型展开测试.....

## 选取的医疗大模型测试平台

### 数量

8个 (其中7个国内平台以及GPT4平台)

## 选取的医疗大模型测试平台

### 类型

综合类大模型平台、垂类医疗大模型平台

## 测试题目

国家执业医师资格考试

## 测试方式

API接口测试/网页对话界面输入测试题目

## 测试时间

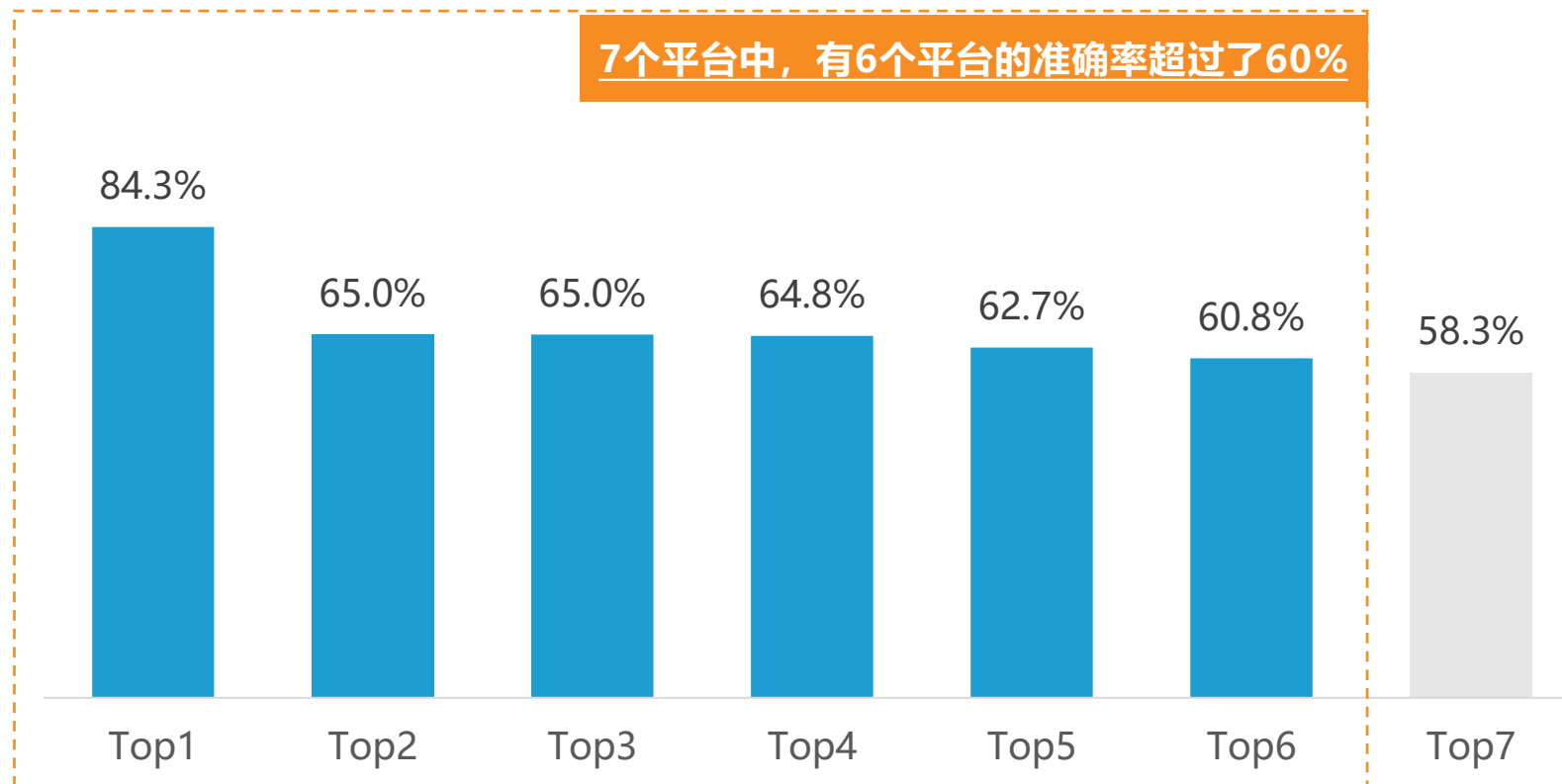
2023年10月-2023年11月

◆ **七个平台中，有六个平台都通过了执业医师综合笔试的考试。**根据国家执业医师资格考试要求，执业医师考试有600道选择题，每道题1分，合格分数线为360分，即准确率超过60%为通过。在本次测试结果中，有六个平台的准确率都超过了60%，而公开数据显示执业医师综合笔试通过率在20%-25%之间，由此可见，医疗大模型平台可能具有比一般考生更高的专业水平或者更好的准备情况。

### 亿欧智库：国内七个大模型平台关于执业医师题目的首次测试结果

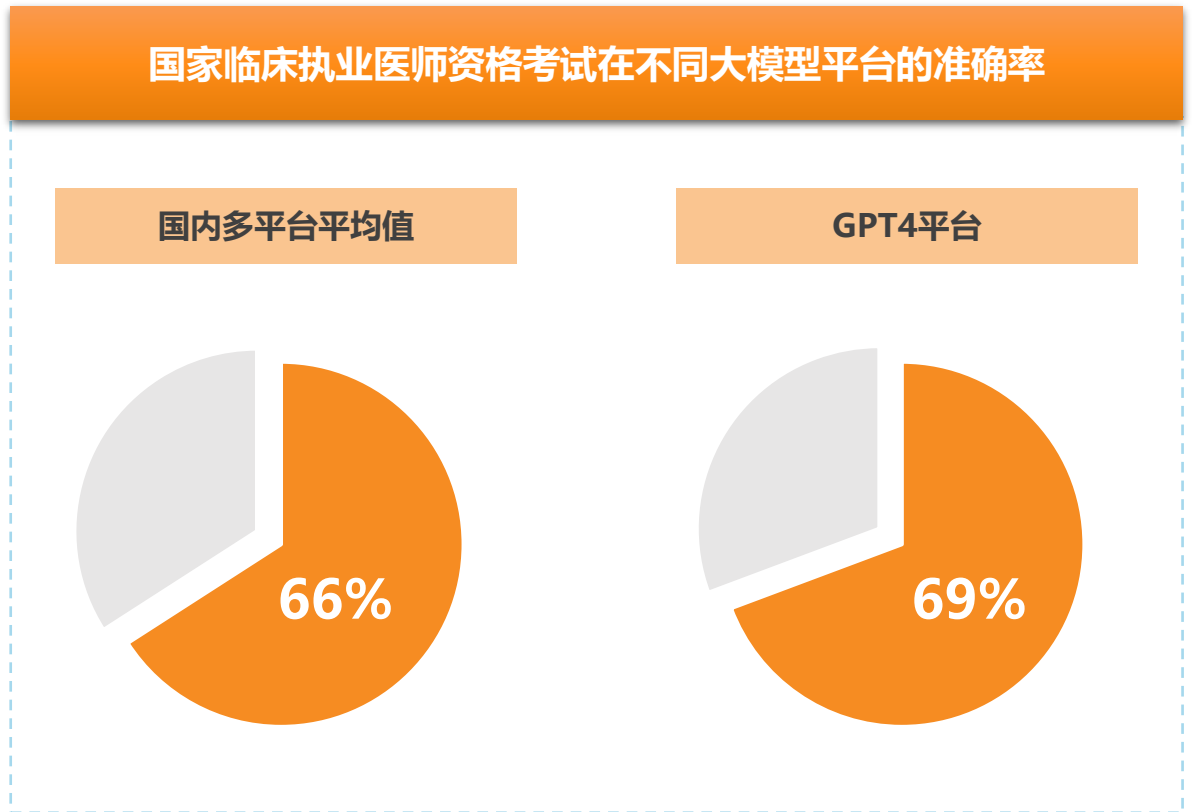
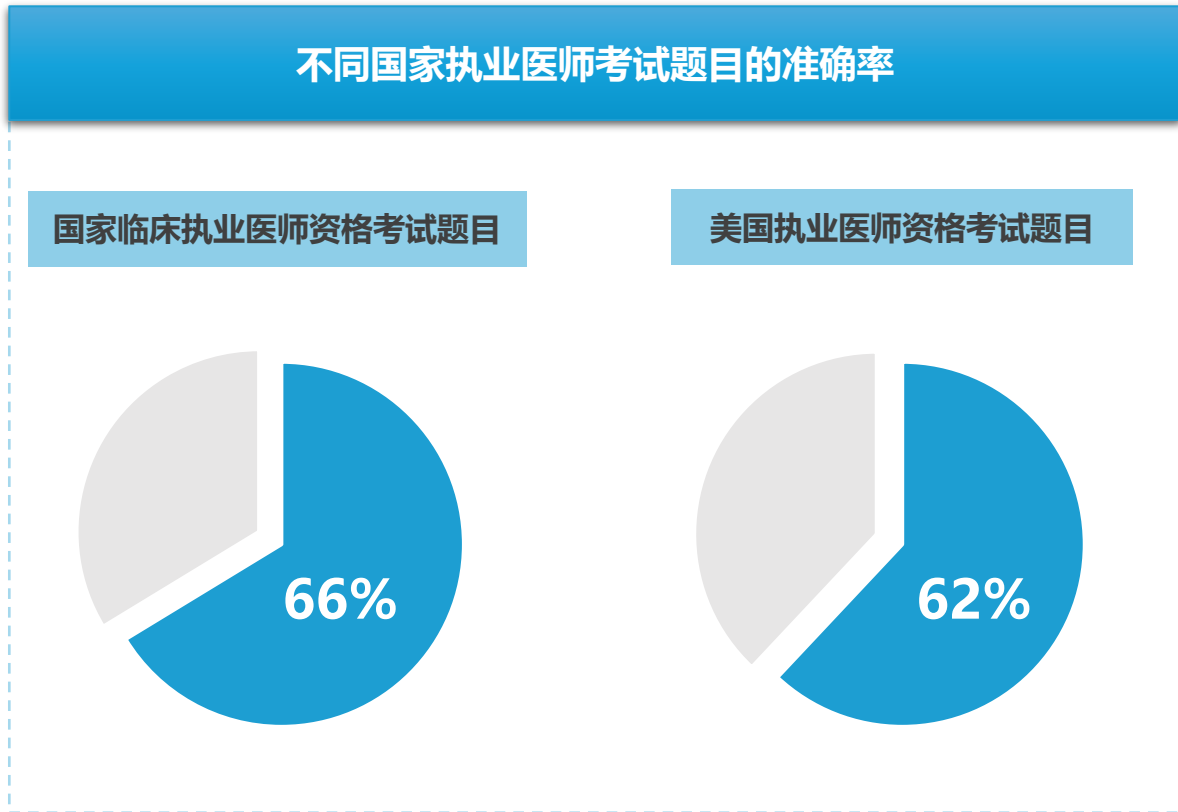
#### 大模型平台在测试执业医师题目过程中，存在不稳定性

- 在各平台测试过程得出结果后，我们将部分错误题目在同一平台上**再次输入，发现两次返回的答案存在差异**，包括第一次答案正确、第二次错误或反之的情况。
- 由于大模型平台测试返回答案存在不稳定性，所以测试结果选取了首次返回的答案用以计算准确性。
- 大模型平台测试存在不稳定性原因可能是由于模型过载、数据问题、网络波动、系统兼容性、动态环境变化以及配置错误等等



# 医疗大模型在做不同国家执业医师考试题目时，准确率差异较小，GPT4的表现略优于大部分国内医疗大模型平台

- ◆ **医疗大模型在做不同国家执业医师考试题目时，准确率差异较小。** 《Performance of Multimodal GPT-4V on USMLE with Image: Potential for Imaging Diagnostic Support with Explanations》一文中指出，GPT-4V，在回答带有图像的美国医学执照考试问题时，取得了62%的准确率。由于带有图像的考试问题难度较大，文字类分析难度较小，但考虑到中文分析理解难度较大，因此二者的准确率相当。
- ◆ **GPT4的表现优于大部分国内医疗大模型平台。** GPT4的准确率为69%，国内平台平均值为66%。

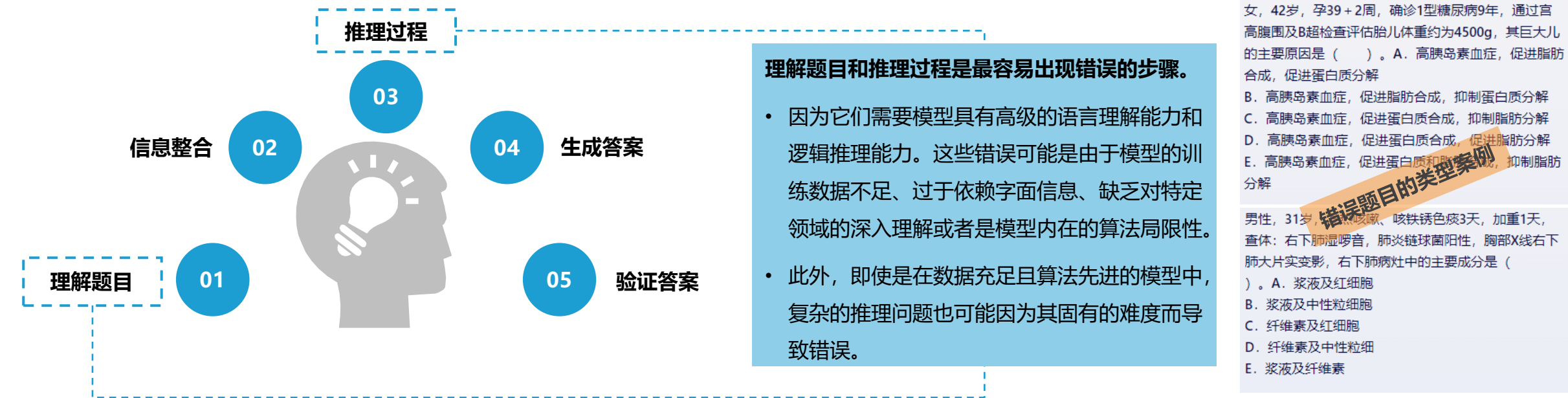




# 医疗大模型对于中文执业医师考试需要分析的案例题目，错误率较高

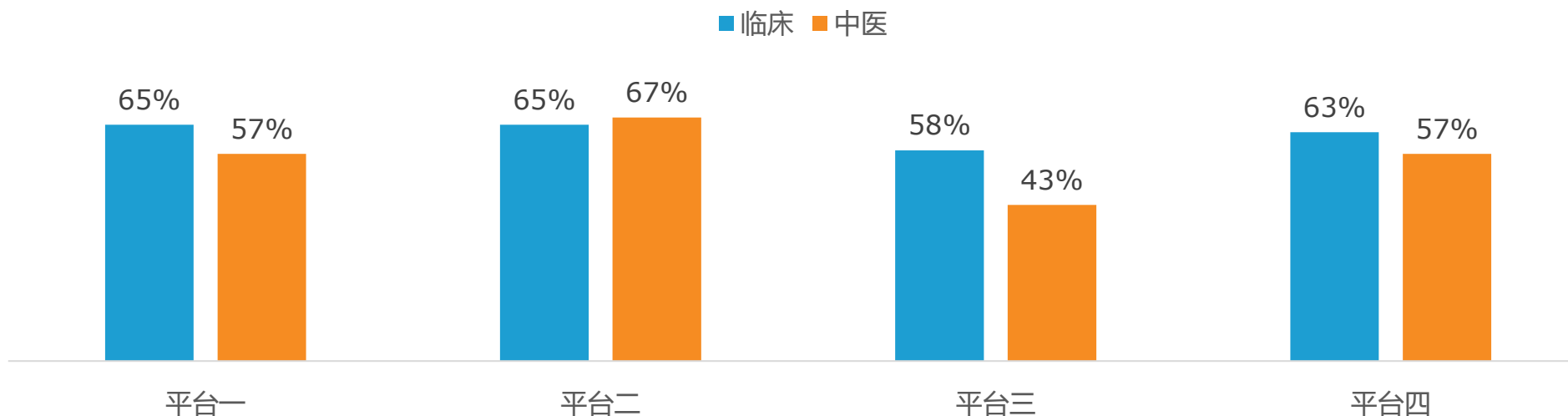
- ◆ 在分析八个平台的返回的答案后，我们发现对于场景类题目，医疗健康AI大模型产品回答的错误率较高。如在题目阐述了性别、年龄、患者症状后，需要给出诊断、或下一步治疗的建议时，往往会出现分析错误的情况。根据回答情况分析，目前错误题目中，34%的题目都是场景相关的题目（即题干中给出详细信息，需要判断）。
- ◆ 因此，亿欧认为，虽然整套试题的准确率较高，但在面临医疗实际行为时，**医疗健康AI大模型仍然无法对复杂的医疗信息给出专业的判断。其使用场景更适合用于院外场景，如患者在挂号前的问诊（具体需要挂哪一个科室）等。**
- ◆ 除此之外，我们发现在中文执业医师考试题目中，类似“选择不正确答案”的题目，错误率较高，也反映出了医疗健康AI大模型在进行逻辑分析时存在漏洞。

## 亿欧智库：国内七个大模型平台关于执业医师题目的首次测试结果



- ◆ 根据测试结果显示，部分医疗大模型在回答中医执业医师题目的准确率略低于临床执业医师的题目。
- ◆ 医疗健康AI大模型在处理中医相关问题时存在以下几个方面的难点：1) **文化和语境的理解**：中医的理论包含了大量的象征性语言和概念（如阴阳、五行等）。医疗大模型可能难以准确理解和解释这些概念的深层含义。2) **诊疗方法的差异**：中医诊断依赖于医师对患者的观察、听诊、问诊和切脉等手段，这些方法与西医的诊断方式（如实验室检测、影像学检查）有很大差异。大模型可能难以模拟中医的诊疗过程，因为这需要对非标准化信息的高度解读能力。3) **治疗个体化**：中医强调根据个体的体质和病情来调整治疗方案，这种个体化的治疗方法要求医师具有丰富的经验，而这些是大模型难以完全复制的。4) **跨学科整合的复杂性**：将中医与现代医学知识结合起来是一个跨学科的挑战，需要深入理解两种体系的原理和方法。大模型可能在没有足够跨学科训练数据的情况下难以实现有效的整合。

亿欧智库：国内四个大模型平台关于部分中医执业医师题目的首次测试结果





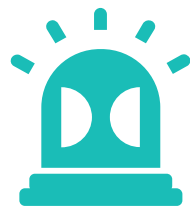
## 结论一：医疗健康AI大模型在执业医师资格考试中表现优异，甚至超过真实考生。

- **医疗健康AI大模型在事实型题目有突出的优势。** 医疗健康AI大模型模型训练于海量的医学文献、指南、教科书和临床案例，且医疗健康AI大模型信息检索能力极强，能迅速准确地提供答案，特别是在选择题和需要直接回忆事实的题目上表现突出。
- **医疗健康AI大模型在答题过程中不会收到情绪干扰。** 医疗健康AI大模型提供的答案通常基于医学指南和证据，因此更加标准化，而人类考生可能受个人经验和偏见影响。



## 结论二：在面对非标准化或模糊的临床情况时，AI无法像经验丰富的医生作出最佳判断。

- **在处理非标准化或模糊的临床情况时无法像经验丰富的医生那样作出最佳判断，主要是因为它缺乏医生的临床直觉、深度的情境理解、灵活的适应性和人际沟通能力。** 在诊疗尤其是在中医的诊疗场景中，医学中常常存在不确定性和模糊性，医生能够基于经验和临床知识处理这些不确定性，做出最合适的判断。
- AI也难以处理医疗决策中的伦理道德问题，以及在面对不确定性时运用创造性思维。



## 结论三：稳定性不足确实是医疗健康AI大模型在实际应用中面临的一个重要障碍。

- 如果AI系统在不同的临床环境和患者群体中表现出波动或不可预测的结果，它们的可靠性和信任度就会受到质疑。**这可能导致误诊、治疗延误或不必要的治疗，增加患者的风险和医疗成本。**
- **因此，确保医疗健康AI系统的稳定性和可泛化性是推广其临床应用的关键。** 这需要通过广泛的验证、持续的监测和严格的临床测试来实现，以确保系统在各种条件下都能提供准确和一致的支持。



## 趋势一

### 应用场景拓展：

- 从诊断辅助向治疗规划和患者监护扩展，例如在远程医疗和慢病管理中的应用。
- AI在药物发现和设计中的作用将进一步增强，缩短药物研发周期。
- AI在公共卫生和疾病预防领域的应用，如通过大数据分析预测突发性的疾病爆发。



## 趋势二

### AI相关法案加速建立：

- 相关部门建立标准化的评估流程，对医疗健康AI应用的安全性和有效性进行验证。
- 制定专门的法律法规，明确医疗健康AI的责任归属和使用规范。
- 相关部门加强对医疗健康AI系统的伦理审查，确保其决策过程公正、透明。



## 趋势三

### 生成式AI加快了人机交互协助：

- 生成式AI增强医疗人员对AI工具的信任和依赖性，通过培训和教育提高他们的AI素养。
- 各科技企业将开发更为直观的用户界面，使医疗人员能够更容易地与AI系统交互。
- 各科技企业将研究人机交互的最佳实践，确保医疗健康AI辅助下的决策过程既高效又人性化。

## ◆ 团队介绍:

亿欧智库 (EO Intelligence) 是亿欧旗下的研究与咨询机构。为全球企业和政府决策者提供行业研究、投资分析和创新咨询服务。亿欧智库对前沿领域保持着敏锐的洞察，具有独创的方法论和模型，服务能力和质量获得客户的广泛认可。

亿欧智库长期深耕新科技、消费、大健康、汽车出行、产业/工业、金融、碳中和等领域，旗下近100名分析师均毕业于名校，绝大多数具有丰富的从业经验；亿欧智库是中国极少数能同时生产中英文深度分析和专业报告的机构，分析师的研究成果和洞察经常被全球顶级媒体采访和引用。

以专业为本，借助亿欧网和亿欧国际网站的传播优势，亿欧智库的研究成果在影响力上往往数倍于同行。同时，亿欧内部拥有一个由数万名科技和产业高端专家构成的资源库，使亿欧智库的研究和咨询有强大支撑，更具洞察性和落地性。

## ◆ 报告作者:



王思晗

亿欧大健康研究副总监  
Email: wangsihan@iyiou.com

## ◆ 报告审核:



高昂

亿欧董事总经理，亿欧大健康总裁  
Email: gaoang@iyiou.com



王辉

亿欧智库副院长  
Email: wanghui@iyiou.com

## ◆ 版权声明:

本报告所采用的数据均来自合规渠道，分析逻辑基于智库的专业理解，清晰准确地反映了作者的研究观点。本报告仅在相关法律许可的情况下发放，并仅为提供信息而发放，概不构成任何广告。在任何情况下，本报告中的信息或所表述的意见均不构成对任何人的投资建议。本报告的信息来源于已公开的资料，亿欧智库对该等信息的准确性、完整性或可靠性作尽可能的追求但不作任何保证。本报告所载的资料、意见及推测仅反映亿欧智库于发布本报告当日之前的判断，在不同时期，亿欧智库可发出与本报告所载资料、意见及推测不一致的报告。亿欧智库不保证本报告所含信息保持在最新状态。同时，亿欧智库对本报告所含信息可在不发出通知的情形下做出修改，读者可自行关注相应的更新或修改。

本报告版权归属于亿欧智库，欢迎因研究需要引用本报告内容，引用时需注明出处为“亿欧智库”。对于未注明来源的引用、盗用、篡改以及其他侵犯亿欧智库著作权的商业行为，亿欧智库将保留追究其法律责任的权利。

## ◆ 关于我们:

亿欧是一家专注科技+产业+投资的信息平台和智库；成立于2014年2月，总部位于北京，在上海、深圳、南京、纽约设有分公司。亿欧立足中国、影响全球，用户/客户覆盖超过50个国家或地区。

亿欧旗下的产品和服务包括：信息平台亿欧网 (iyiou.com)、亿欧国际站 (EqualOcean.com)、研究和咨询服务亿欧智库 (EO Intelligence)，产业和投融资数据产品亿欧数据 (EO Data)；行业垂直子公司亿欧大健康 (EO Healthcare) 和亿欧汽车 (EO Auto) 等。

◆ 基于自身的研究和咨询能力，同时借助亿欧网和亿欧国际网站的传播优势；亿欧为创业公司、大型企业、政府机构、机构投资者等客户类型提供有针对性的服务。

## ◆ 创业公司

亿欧旗下的亿欧网和亿欧国际站是创业创新领域的知名信息平台，是各类VC机构、产业基金、创业者和政府产业部门重点关注的平台。创业公司被亿欧网和亿欧国际站报道后，能获得巨大的品牌曝光，有利于降低融资过程中的解释成本；同时，对于吸引上下游合作伙伴及招募人才有积极作用。对于优质的创业公司，还可以作为案例纳入亿欧智库的相关报告，树立权威的行业地位。

## ◆ 大型企业

凭借对科技+产业+投资的深刻理解，亿欧除了为一些大型企业提供品牌服务外，更多地基于自身的研究能力和第三方视角，为大型企业提供行业研究、用户研究、投资分析和创新咨询等服务。同时，亿欧有实时更新的产业数据库和广泛的链接能力，能为大型企业进行产品落地和布局生态提供支持。

## ◆ 政府机构

针对政府类客户，亿欧提供四类服务：一是针对政府重点关注的领域提供产业情报，梳理特定产业在国内外的动态和前沿趋势，为相关政府领导提供智库外脑。二是根据政府的要求，组织相关产业的代表性企业和政府机构沟通交流，探讨合作机会；三是针对政府机构和旗下的产业园区，提供有针对性的产业培训，提升行业认知、提高招商和服务域内企业的水平；四是辅助政府机构做产业规划。

## ◆ 机构投资者

亿欧除了有强大的分析师团队外，另外有一个超过15000名专家的资源库；能为机构投资者提供专家咨询、和标的调研服务，减少投资过程中的信息不对称，做出正确的投资决策。

## ◆ 欢迎合作需求方联系我们，一起携手进步；电话 010-53321289，邮箱 hezuo@iyiou.com





扫码关注亿欧智库  
查看更多研究报告



扫码添加小助手  
加入行业交流群



网址: <https://www.iyiou.com/research>

邮箱: [hezuo@iyiou.com](mailto:hezuo@iyiou.com)

电话: 010-53321289

北京: 北京市朝阳区关庄路2号院中关村科技服务大厦C座4层 | 上海: 上海市闵行区申昆路1999号4幢806

深圳: 广东省深圳市南山区华润置地大厦 C 座 6 层 | 纽约: 4 World Trade Center, 29th Floor-Office 67, 150 Greenwich St, New York, NY 10006