

# IBM 的 AI 医疗走进死胡同：想用 NLP 解决医学问题，连 Bengio 也不看好

人工智能 2019-04-16

从 2011 年开始，蓝色巨人 IBM 在 AI 医疗上押上了重注，甚至为其 AI 部门建立了一个华丽的总部。这一切都是为了在 AI 浪潮中保持自己的领先定位。



然而 IBM Watson 的医疗技术已经使用了 8 年，却从未给人留下深刻印象。

IEEE Spectrum 高级副主编 Eliza Strick 认为 IBM Watson 医疗雷声大雨点小，前期给了很高承诺，如今却远远没有达成目标。

在 IBM 的设想中，Watson 的记忆库掌握着各种罕见疾病的知识，可以在几秒内就解决一个棘手的病例。如果 Watson 能够将这种即时专业知识带到世界各地的医院和诊所，人工智能似乎可以减少诊断错误。

但是 Watson 的成果只限于研究领域。IBM 试图将 Watson 应用于医学界的最大挑战之一癌症时，遇到了机器学习方式和医生工作方式之间的根本不匹配。

8 年承诺未能兑现

2011 年，IBM Watson 在智力竞赛电视节目《危险边缘》中战胜两位人类冠军，随后 IBM 宣布了自己雄心勃勃的计划：未来让 Watson 成为一名 AI 医生。IBM 还承诺在未来 18 到 24 个月的时间里推出首批医疗健康产品。

实际上，IBM 当时宣布的项目并没有产生任何商业产品。

此后 8 年中，IBM 高调宣扬自己在 AI 医疗技术上的努力，但是与其他医院、高校合作的 AI 医疗项目大多数已经失败。

据 IEEE Spectrum 统计，从 2011 年开始，IBM Watson 与其他机构合作的 25 个具有代表性的项目中，仅有 5 个合作项目推出了 AI 医疗产品。

DATE	IBM PARTNER	PROJECT	CURRENT STATUS
<b>2011</b> Feb.	<u>Nuance Communications</u>	<u>Diagnostic tool and clinical-decision support tools</u>	No tools in use
Sept.	<u>WellPoint (now Anthem)</u>	<u>Clinical-decision support tools</u>	No tools in use
<b>2012</b> March	<u>Memorial Sloan Kettering Cancer Center</u>	<u>Clinical-decision support tool for cancer</u>	Watson for Oncology
Oct.	<u>Cleveland Clinic</u>	<u>Training tool for medical students; clinical-decision support tool</u>	No tools in use
<b>2013</b> Oct.	<u>MD Anderson Cancer Center</u>	<u>Clinical-decision support tool for cancer</u>	No tool in use
<b>2014</b> March	<u>New York Genome Center</u>	<u>Genomic-analysis tool for brain cancer</u>	No tool in use
June	<u>GenieMD</u>	<u>Consumer app for personalized medical advice</u>	No app available
Sept.	<u>Mayo Clinic</u>	<u>Clinical-trial matching tool</u>	Watson for Clinical Trial Matching
<b>2015</b> April	<u>Johnson &amp; Johnson</u>	<u>Consumer app for pre- and postoperation coaching; consumer app for managing chronic conditions</u>	No apps available
April	<u>Medtronic</u>	<u>Consumer app for personalized diabetes management</u>	Sugar.IQ app
May	<u>Epic</u>	<u>Clinical-decision support tool</u>	No tool in use
May	<u>University of North Carolina, others</u>	<u>Genomic-analysis tool for cancer</u>	Watson for Genomics
July	<u>CVS Health</u>	<u>Care-management tool for chronic conditions</u>	No tool in use
Sept.	<u>Teva Pharmaceuticals</u>	<u>Drug-development tool; consumer app for managing chronic conditions</u>	No tool in use; no app available
Sept.	<u>Boston Children's Hospital</u>	<u>Clinical-decision support tool for rare pediatric diseases</u>	No tool in use
Dec.	<u>Nutrino</u>	<u>Consumer app for personalized nutrition advice during pregnancy</u>	No app available
Dec.	<u>Novo Nordisk</u>	<u>Consumer app for diabetes management</u>	No app available
<b>2016</b> Jan.	<u>Under Armour</u>	<u>Consumer app for personalized athletic coaching</u>	No app available
Feb.	<u>American Heart Association</u>	<u>Consumer app for workplace health</u>	No app available
April	<u>American Cancer Society</u>	<u>Consumer app for personalized guidance during cancer treatment</u>	No app available
June	<u>American Diabetes Association</u>	<u>Consumer app for personalized diabetes management</u>	No app available
Oct.	<u>Quest Diagnostics</u>	<u>Genomic-analysis tool for cancer</u>	Watson for Genomics from Quest Diagnostics
Nov.	<u>Celgene Corp.</u>	<u>Drug-safety analysis tool</u>	No tool in use
<b>2017</b> May	<u>MAP Health Management</u>	<u>Relapse-prediction tool for substance abuse</u>	No tool in use

IBM 还试图通过收购来加强自身 ,但内部人士表示 ,收购的公司并没有起太大的作用。从 IBM 的 WatsonHealth 部门出来的产品与曾经设想的 AI 医生完全不同 :它更像是可以执行某些日常任务的 AI 助手。

加州大学旧金山分校医学系主任 RobertWachter 说 : IBM 遇到了麻烦。他们把营销放在第一位 ,产品放在第二位 ,吊足了人们的胃口 ,但是项目上马后 ,又发现这是一个非常难解决的问题。

**NLP 短期解决不了问题**

目前 ,美国监管机构只批准了少数 AI 工具用于医院。这些开创性的产品主要是在视觉领域 ,而 IBM 却一直致力于用 NLP 解决医疗问题。

其他 AI 医疗公司在研究用计算机视觉分析图像 ,如 X 光片和视网膜扫描。但是 IBM 没有这样分析医学影像的产品 ,尽管他们在该领域有一个活跃的研究项目。

除去医疗影像 ,即便是今天最好的 AI 也难以理解复杂的医疗信息。用软件对人类医生专业知识进行编码是一件非常困难的事情。正如全世界所看到的 ,IBM 已经在市场上学到了这些痛苦的教训。

虽然 IBM 并没有放弃目标 ,但它的失败已经向技术专家和医生们表明 ,创建 AI 医生是多么困难。



2011年 Watson 在知识问答上的胜利，证明了它在 NLP 方面的卓越表现。为了玩这个游戏，它必须解析复杂的文字游戏线索，搜索大量的文本数据库，找到并确定最好的答案。

“看起来 Watson 几乎可以理解语言的含义，而不仅仅是识别单词的模式。”在 2011 年曾担任 IBM 研究院首席医学科学家 Martin Kohn 说。

其实早在参加《危险边缘》之前，IBM 就已经考虑过 AI 医疗的可能性。

大量的患者数据看起来非常适合用于 AI 医疗研究，特别是当医院和医生开始使用电子健康记录时。

虽然其中一些数据可以很容易地被机器消化，例如实验室结果和生命体征测量数据，但还有大部分是“非结构化”信息，例如医生笔记和出院记录。这类叙述性文本占到患者记录约 80% 的内容。

Kohn 认为，Watson 强大的 NLP 能力可以转变为医学理论。Watson 可以阅读患者的健康记录以及医学文献的全部内容，包括教科书、同行评审期刊文章、批准药物清单等。通过访问所有这些数据，Watson 可能会成为一名超级医生，能够辨别出人类无法看到的模式。

哥伦比亚大学医学系教授 Herbert Chase 曾与 IBM 展开合作，希望 Watson 能成为一种“临床诊断支持”工具，他们一起研究了一种诊断工具的原型。但是双方从 2014 年分道扬镳，原因是 Chase 教授对 Watson 进展缓慢感到失望。

Chase 教授只是 IBM Watson 的众多合作者之一，另一名合作者马里兰大学放射学教授认为，AI 工具在 10 年内对医生来说是不可或缺的，但他并不相信 IBM 能实现它们。

他说：“我不认为 IBM 处于 AI 技术的最前沿，最令人兴奋的事情发生在苹果、谷歌、亚马逊这些公司。”

2014 年，Kohn 离开了 IBM，他说：“仅仅证明你拥有强大的技术是不够的，还需要向别人证明你能做出有用的东西。”

Kohn 一直在等待医学期刊上的论文证明 AI 可以改善患者的治疗效果并节省医疗系统的费用。到目前为止，这类文章发表的很少，Watson 也没有任何结果。

### IBM 首次涉足 AI 医疗领域

在试图将 AI 引入医疗时，IBM 面临着巨大的挑战，它在许多计算领域已经落后于谷歌、苹果等巨头。IBM 需要一些重要的东西才能保证地位。为此，IBM 在 AI 医疗领域投入大量资金。

2014 年，IBM 向 Watson 部门投资 10 亿美元。2015 年，IBM 宣布成立一个特殊的 WatsonHealth 部门，到 2016 年中期，WatsonHealth 已经收购了四家健康数据公司，总花费约 40 亿美元。

IBM 似乎有了让 AI 进入医疗健康领域所需的技术、资源和承诺。“将 AI 用于医疗健康是一项艰巨的任务，这是一项挑战，但我们正在这样做。”IBM 认知解决方案和 IBM 研究院高级副总裁 John E. Kelly 说。从 2011 年开始，Kelly 就负责指导 Watson 研究，现在他还负责对 WatsonHealth 的直接监督。

他说公司在需要的时候已经转向：“我们正在继续学习，所以我们的产品会随着我们的学习而改变。”

IBM 的医疗诊断工具就是一个例子，它并没有上市，IBM 医疗保健和生命科学研究副总裁 AjayRoyyuru 说：“诊断不是我们要进入的领域，这是医疗专家们可以做得很好的地方，无论你用 AI 做得多好，都不会取代人类专家。”

为了找到 AI 医疗的商业案例，IBM 开展了一系列令人眼花缭乱的项目，针对医疗保健系统中不同参与者：医生、行政人员、保险公司和患者。

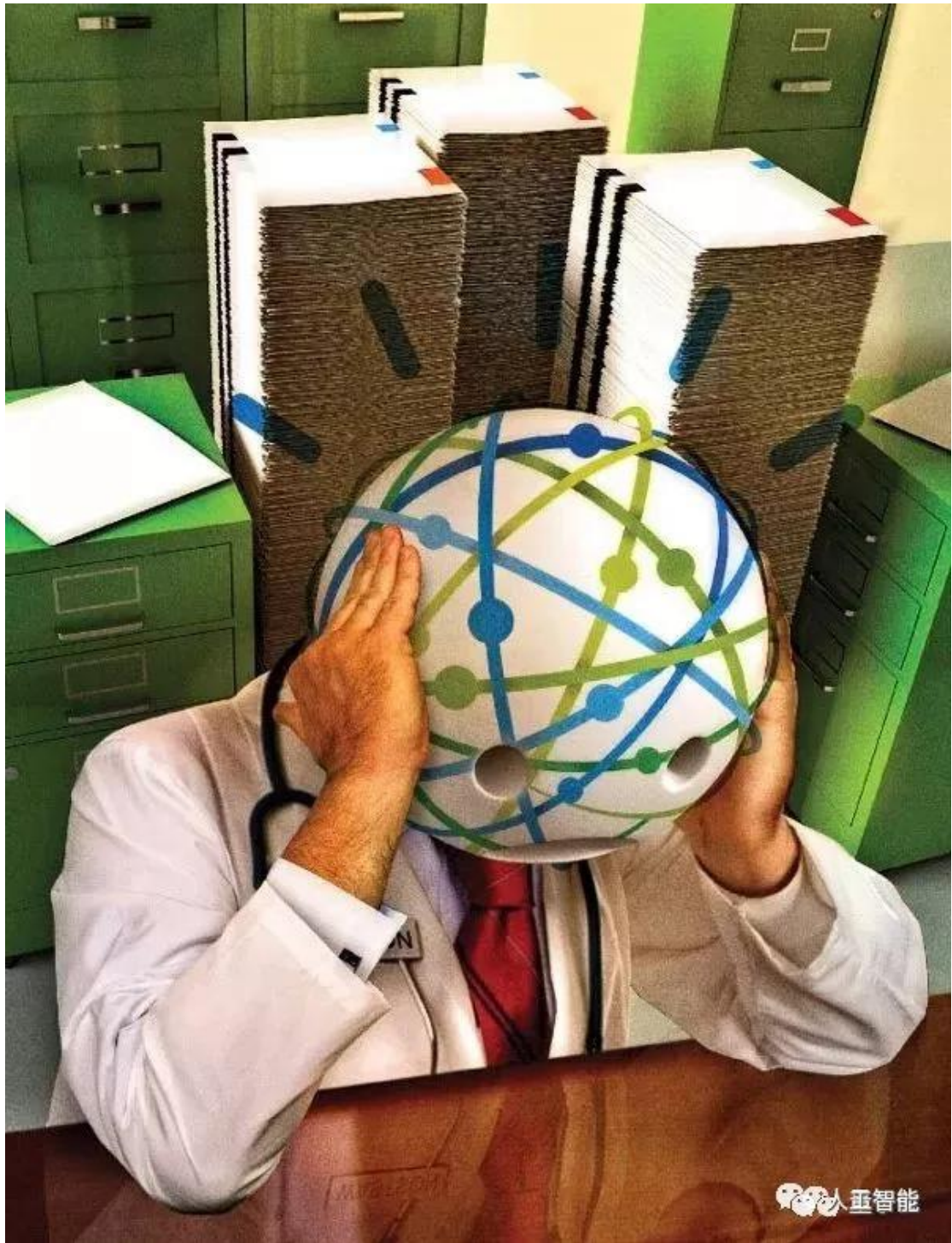
在许多种尝试中，IBM 努力让 Watson 的 NLP 去理解医疗文本。但蒙特利尔大学计算机教授 YoshuaBengio 说，在医学文本文件中，人工智能系统无法理解其模糊性，也无法了解人类医生注意到的微妙线索。到目前为止，没有人工智能可以达到人类医生的理解和洞察力。

IBM 对癌症的研究是这家公司遇到的主要挑战之一。一位肺癌专家 MarkKris 表示：“我认为没有人会知道这（实现癌症诊断）会花多长时间。”他所在的研究机构早在 2012 年就和 IBM Watson 展开了合作。

Kris 和其他医生在 2015 年训练了一个人工智能系统，该系统成为了 Watson 肿瘤产品。德克萨斯大学 MD 安德森癌症中心的医生与 IBM 合作创建了一个名为 Watson 肿瘤顾问。白血病部门测试了该工具，但它从未成为商业产品。

Watson 肿瘤顾问遭受了多方批评。有文章指出 Watson 肿瘤是无用甚至会提供危险的建议。Kris 说他经常听到批评说该产品不是“真正的人工智能”。IBM 与 MD 安德森癌症中心的合作项目花费了 6200 万美元，最后以失败告终而被取消。





该项目揭示了机器学习的承诺与医疗保健的现实根本不匹配，在“真正的 AI”与当今医生对功能性产品的要求之间存在着差距。

Watson 肿瘤通过获取大量癌症患者健康记录和医学文献来学习。IBM 希望 Watson 凭借其强大的计算能力，将检查这些记录中的数百个变量，发现人类无法看到的模式。对于肿瘤学家来说，这听起来像癌症护理方面的潜在突破。对于 IBM 来说，它听起来就像一个伟大的产品。

Watson 很快就学会了如何浏览临床研究的文章并确定基本结果。但它不会像医生那样阅读文章。医生是从文章中提取信息，用来改变他们的治疗方法。

而 Watson 的思想是基于统计数据，所以它能做的就是收集有关主要结果的统计数据。“但医生不这样做。”

例如，在 2018 年，FDA 批准了一种新癌症药物，它对所有表现出特定基因突变的肿瘤都有效。人类医生会建议每个肺癌患者都接受这种基因检测，但 Watson 不会根据有 55 名患者中 4 名肺癌患者改变结论。

与现有医疗体系标准不符

Watson 无法从医疗文献中新的病历事件提取出自己的见解，这只是它失败的第一步。研究人员还发现，它也无法像预期的那样从患者的电子健康记录中挖掘信息。

研究人员让 Watson 研究白血病患者健康记录。在一份 2018 年发表于 TheOncologist 的论文中，该团队报告说，基于 Watson 的肿瘤学专家顾问在从医疗记录中的文本文档中提取信息方面取得了不同程度的成功。在处理诊断等明确概念时，准确度从 90% 到 96% 不等，但对依赖于时间的信息（如治疗时间线），准确度仅为 63% 至 65%。

不仅如此，IBM 实现 AI 超级医生的梦想还遭遇了另一重打击：研究人员意识到，Watson 无法将新患者与之前发现隐藏模式的癌症患者进行比较。IBM 本来希望 AI 能够模仿他们的专家肿瘤学家的能力，利用之前的结果和经验，为新患者制定治疗策略。

但医疗保健系统目前的标准并不鼓励这种现实世界的学习。AI 系统的建议基于它在医疗记录中发现的模式，但是如果没有科学研究的严格控制，这种发现只会被认为是相关性，而不是因果关系。

IBM 认为医疗保健的标准必须改变，才能让 AI 实现其全部潜力并转化为医学。

### 肿瘤专家不信任

根据传闻报道，IBM 在美国寻找其 Watson 肿瘤产品的买家时遇到了麻烦。一些肿瘤学家说他们相信自己的判断，并不需要沃森告诉他们该做什么。有人说 Watson 只是建议他们很清楚的标准治疗方法。

在过去的几年里，与 IBM 合作的医院已开始发表关于 Watson 肿瘤相关经验的研究。

在印度，Manipal 综合癌症中心的医生对 638 例乳腺癌病例中，Watson 与专家治疗建议的一致率为 73%。Watson 在韩国 Gachon 大学 Gil 医疗中心表现更差，该医院为 656 名结肠癌患者提供的最佳建议，只有 49% 与专家相匹配。

WatsonHealth 也有一些成功案例。IBM 与北卡罗来纳大学的合作伙伴在 2017 年发表了第一篇关于 WatsonforGenomics 有效性的论文中，参加该研究的 32% 的癌症患者，Watson 发现了人类从未发现的潜在重要突变。

WatsonHealth 的故事是一个警示，每个人都喜欢雄心壮志，但是没人希望被放鸽子。

**- END -**