



# 2023 医疗大数据 白皮书

数据资产化  
大数据新基建  
开启千亿市场序幕

*BIG DATA*

## 前言

过去十年，国内医疗大数据产业从零开始发展，如今已然初具规模。但受地域政策、数据伦理等因素影响，整个产业一度在 2019 年抵达巅峰，随后陷入停滞，直至今年才又重新复苏。

推动产业回到“高速公路”的动力主要分为两个层面。政策层面，《“十四五”大数据产业发展规划》出台，数据被重新界定为“生产要素”，重要性直接升至国家战略高度；“数据二十条”的发布与国家数据局的建立，为大数据产业基础制度体系指明建设方向，确立指导制度建设责任主体。

技术层面，新兴的大语言模型、生成式 AI 借助底层逻辑的重构带来了一批类似 Chat GPT 的“杀手级应用”，反哺上游大数据产业发展。该趋势下，医疗数据持有方开始进行主动改革，规模化开展医疗数据治理，为未来可能的数据资产化做好准备。

多重因素复合之下，医疗大数据产业迎来重大发展机遇。为了厘清医疗大数据的发展现状及未来价值，蛋壳研究院此份报告，全面梳理医疗大数据产业，绘制医疗大数据的产业前景。

## 核心观点

一、因担忧数据因素、数据安全方面的风险，医疗大数据产业在 2020 年前一直重基建而轻应用，但自 2020 年开始国家政策风口开始转向，“数据二十条”发布、国家数据局建立，均在推动医疗大数据朝着标准化、产业化的方向发展，进而建立完善的交易体系。

二、伴随信息技术和网络技术的跨越式发展，现代医院的运营特征表现为医疗业务的智能化与应用部署的敏捷化，使得医院业务产生的信息朝着复杂化、专业化、海量化的方向发展，并对各系统之间的互联互通提出更高层级的要求。因此，部分医院开始转变信息化建设思路，借助云技术打造新一代医院数据中心，综合利用各类数据服务临床、决策和科研过程，提高医院管理的科学化、规范化、精细化水平，由此创造了一个千亿级的新兴医疗 IT 市场。

三、尽管基础设施已经初具规模，但多数医院大数据应用开拓情况仍处于低位，三级医院应用数量占比不足 20%，二级医院应用数量占比不足 5%。因此，在大数据应用供不应求的情形下，在找到有效商业模式的前提下开拓大数据应用是医疗 IT 公司面临的重要机遇。

四、大语言模型（LLM）的出现为医疗大数据产业确立了新的增长点。目前医疗垂直大模型使用的数据仍局限于互联网医疗过程中产生的行为数据，未来若能将临床数据规模化纳入其中，国内医院的智能化建设将进一步加速。

# 目录

第一章：数据基础建设即将确定，数据资产化亟待破解 .....	1
1.1 聚焦医院：产生医疗大数据发展问题的原因 .....	2
1.2 聚焦政策：政策频出，医疗大数据宏观环境迎来多轮利好 .....	3
1.3 数据资产化，需打通四重要素 .....	5
第二章：基建扩容，千亿市场的根基 .....	7
2.1 新一代医院数据中心开启千亿市场序幕 .....	9
2.2 影像补位，AI 企业成为基建主角 .....	12
2.3 临床研发发力，科研药研加速数据库建设 .....	13
2.4 物联网、AI 支持的医院智慧后勤 .....	17
2.5 大数据赋能的智慧医学教育 .....	18
2.6 医院大数据中心建设的常见问题探讨 .....	19
第三章：应用创新，数字医疗建设进行时 .....	22
3.1 临床大数据应用 .....	23
3.2 运营大数据应用 .....	26
3.3 院外其他大数据与应用 .....	30
第四章：数据安全，成为应用推广的必要保障 .....	35
4.1 医疗大数据安全现状 .....	36
4.2 大数据安全体系的构建 .....	37
4.3 数据安全的保障逻辑 .....	38
第五章：前沿讨论，大语言模型的构建与数据交易的可能 .....	41
5.1 讨论一：LLM 对于多模态大数据治理的潜在影响 .....	42
5.2 讨论二：应用级医疗大数据的交易可能 .....	43
第六章：企业价值，助推医疗信息化进程的重要力量 .....	47
参考目录： .....	53

# BIG DATA

## 第一章

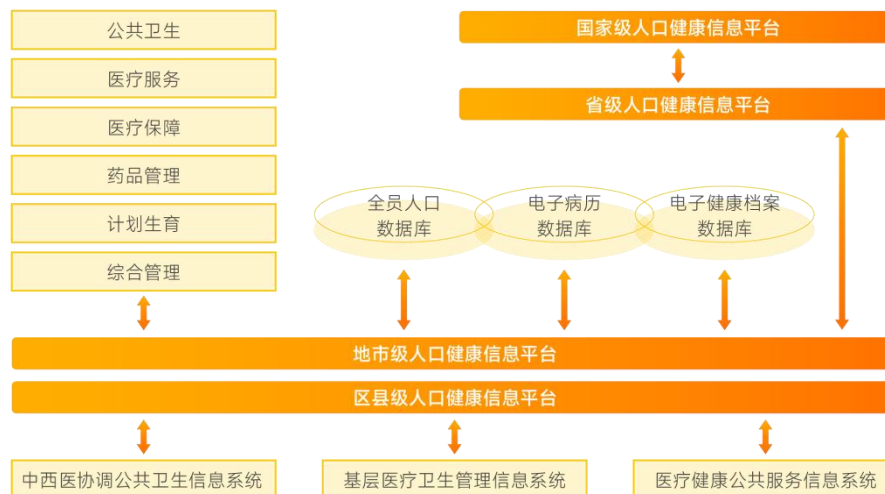
数据基础制度建设

即将确定，数据资产化

亟待破解

2014年，为推动整体化卫生信息系统建设，国家卫计委规划了开创性的“46312”工程。意在搭好医疗大数据基础设施建设的骨架，将碎片化的“大量医疗数据”转化为规范、可用的“医疗大数据”。

图表 1 国家卫计委规划的“46312”工程



数据来源：蛋壳研究院

这是一种“自上而下”的建设模式。沿着这一路径进行推进，框架内的卫生信息平台、基础数据库、融合网络、业务应用、信息安全防护体系、人口健康信息标准体系相继建立，国家层面的数据基础设施基本建设完毕。

不过，完好的顶层架构虽有效推动了区域、医院的大数据相关建设，但却未能全面覆盖整个大数据产业。实际之中，这些建设更多围绕基础设施展开，真正能够传递价值的应用部分相较式微，没有达到等同于基础设施建设的投入规模。

探索这一问题的成因及现阶段的解决之道是本报告的研究目的之一。本章之中，我们将从多个角度分析限制医疗大数据的关键要素，并回答“医疗大数据前景几何”这一问题。

## 1.1 聚焦医院：隐私与认知制约医疗大数据发展

千禧年后的第二个十年，移动医疗、人工智能等前沿技术的发展唤起医疗数据的需求。作为算法、算力、数据三要素中最常见但又最难获取的要素，医疗数据彼时仍以碎片化、非标准化的形态分散于医院各个系统中。为了寻找智能模型所需的养料，大量科技医疗创业公司找到三级医院进行合作，在帮助医院进行数据治理的同时，打造智慧化的临床应用。

医院科室的参与、政策对于信息化建设的强制要求，合力促使医院围绕互联互通、智慧医院等方向开启规模化建设。不少医院开始打造医院大数据中心、科研级大数据平台，完成了医

疗大数据基础设施的构造，也与企业合作开发了不少智慧化的应用。

但在 2019 年中美贸易争端开启后，包含个人私密信息医疗数据成为关注重点之一。由于对此类数据进行治理、集成、应用存在一定泄露风险，医院与企业的合作目的开始转变。

为避免政策风险带来的不确定性，不少医院期望大数据及其研究结果以医院范围为界限展开，使得医疗大数据的研究重心转向医院科研需求。此趋势下，医疗大数据产业转化一定程度减少，医疗大数据行业发展整体放慢。

不过，政治因素并非钳制医疗大数据发展的唯一因素，更需关注的是该类建设投资回报及参与度问题。

对于绝大多数而言医院而言，院内外规范化的 IT 建设是一项难以计量回报的投资，在缺乏合适的工具估算大数据建设的产出时，医院对于相关投资仍然持有保守态度。

此外，要让该项建设发挥价值，医院动用资金支持仅是一部分，更重要的是医院深入了解医疗大数据建设内容，将系统与业务有效融合，才能构造行之有效的大数据体系

从目前来看，尽管存在各类标准推动医疗数据的互通互认、治理应用，但只有完整做好每一类场景全流程数据的收集、清洗、归纳、存储都一系列步骤，才能形成多模态、跨流程、可服务于应用的大数据，真正将医疗数据沉淀下来。

这个建设过程应由所有相关成员的共同参与的过程，目前医院还需提升主体积极性，实现全流程、高参与度的数据治理。只有将医疗大数据的被动应用转变为主动管理，才能真正用好医疗大数据。

## 1.2 聚焦政策：政策频出，医疗大数据宏观环境迎来多轮利好

尽管院内的大数据建设存在诸多问题，但自 2020 开始，政策的加持与技术的推动已在潜移默化中消解这些问题。

2020 年 4 月，中共中央、国务院印发《关于构建更加完善的要素市场化配置体制机制的意见》，将“数据”列为劳动力、土地、资本等之外的第五大生产要素，明确要加快培育数据要素市场，推进政府数据开放共享、提升社会数据资源价值、加强数据资源整合和安全保护，制定出台新一批数据共享责任清单。

2021 年 11 月，《“十四五”大数据产业发展规划》提出了更为精确的总体目标，要求“到 2025 年我国大数据产业测算规模突破 3 万亿元，年均复合增长率保持 25%左右，创新力强、附加值高、自主可控的现代化大数据产业体系基本形成。”

2022 年 12 月，《中共中央 国务院关于构建数据基础制度更好发挥数据要素作用的意见》

(后简称：数据二十条) 对外发布则以**构建基础制度为目标**，从数据产权、流通交易、收益分配、安全治理等四个方面，对制定数据基础制度进行了全面部署，最终构建公平与效率相统一的数据要素按贡献参与分配的制度。

2023 年 3 月，国家数据局组建完毕，中央网络安全和信息化委员会办公室、国家发展和改革委员会共同管理，两大机构将在后续协调推进数据基础制度建设，统筹数据资源整合共享和开发利用，统筹推进数字中国、数字经济、数字社会规划和建设等。

图表 2 国家数据局的管理义务划分



数据来源：蛋壳研究院

上述政策中，《关于构建更加完善的要素市场化配置体制机制的意见》与《“十四五”大数据产业发展规划》为医疗大数据的产业化做好背书，有效提升了医疗机构及相关大数据企业的信心。“数据二十条”、成立国家数据局则聚焦于成立良好的市场制度，完善交易流通所需的市场要素，保障大数据市场公平、高效地开展下去。双管齐下，众多政策突显了国家规范发展数据相关产业的决心。

技术方面的迭代同样推动着医疗大数据产业的发展。2022 年末，大语言模型 (LLM) ChatGPT 的火热带动人们重新审视人工智能的价值，并聚焦于背后支撑应用的技术生成式 AI (Generative AI)。2023 年开始，不少医疗 IT 公司、医疗 AI 公司、互联网医疗公司均已开发出自己的大语言模型，并尝试在医院场景之中开发各项新式 AI 应用。

拆解这一新兴人工智能仍是算法、算力、数据、知识四要素，但对于国内企业而言，算法部分均用的开源模型，算力可以根据需求购置 GPU 实现，知识可以通过向权威知识库购买或达成战略合作得到，唯独数据需要企业与医院达成合作，在脱敏、不出院的情况下训练模型。

政策与技术双向驱动，医疗大数据产业再度火热。如今，更多医院开始参与大数据基础及应用建设，大企业们也嗅到风向，广泛参与其中，为产业注入新的活力。



### 1.3 数据资产化，需打通四重要素

宏观条件完备后，微观层面同样需要跟进，即需要各个医院一改过去满足政策要求、被动建设的模式，主动参与到医疗大数据的相关建设中。

目前，医管部门通过强化医院对于科研相关能力的要求，促使医院主动收集数据推动回顾性试验研究、前瞻性试验研究，已经一定程度推动医院转向主动数据治理。

而要进一步提升这一主动性，转变医院对于医疗大数据的认知态度，则需推动数据资产化，将医疗大数据的治理工程由成本项转变为收入项，让医院医生将数据视作一种极具潜力的资产。

这是一个极具挑战的转变。回顾劳动、土地、资本等生产要素形成的市场，流通是其最大化价值实现的必经之路。对于数据这类新式生产要素，怎样保证流通过程的安全顺畅，怎样保障要素市场的公平稳定等问题，需要多个参与方协力解决。

目前，限制数据流通的核心要素可归纳为以下几点。

- 数据要素服务相关制度亟需完善

数据的交易流通需要数据归属、定价、交易、权益分配、安全等基础制度的保障。当前制度尚未形成统一的保护机制，从国家到省市再到区县，跨委办局、垂管系统的数据共享仍然困难。

国家数据局或在未来解决统筹建设的问题，推动各项数据标准的制定，但从成立到发挥关键作用，尚需时日化解从理论到实践过程中遭遇的种种问题。

- 缺乏有效的数据归属权界定标准

目前尚无法律法规对医疗数据的归属权进行界定。有观点认为医疗大数据反映的是个人的健康状况，理应属于患者个人；有观点认为医疗大数据是由医院采集、录入才能产生的，存储和保存也在医疗机构，理应属于医疗机构；还有观点认为医疗数据的所有权在于患者个人、控制权在于医院、管理权在于政府，第三方机构需借助政府支持和医院配合方能对其进行商业化开发和利用。

一方面，模糊的数据归属权可能导致利益分配纠纷，进而提高数据流通成本。因此，产业必须及早明确数据归属权并建立合理的分配制度，保障数据流通的公平公正公开。

另一方面，由于数据易复制的特性，数据提供给其他实体后，两者是价值关系，而不是一次性收益。因此，需要法律法规确保供应商获得应得、持续的收益流。

- 市场机制尚无法满足流通交易需求

数据的需求方和数据的拥有方当前没有合适的发现对接机制，市场没有形成上下游的关系。专利权和著作权有国家知识产权保护，数据权益当前缺少相应法律法规的保护。

虽然国内已有不少大数据交易所开始运营，但数据要素流通市场整体形式仍较为单一，通过交易所挂牌数据较少，从量和质上都无法满足数据市场的需求。相较之下，大量的场外数据交易市场活跃，但缺乏有效监管和安全保障。

医疗方面，已有不少交易所将为“医疗卫生”数据设置交易品类，但绝大部分交易所该品类下并无供应商品，仅贵州大数据交易所上架了一款“儿童构音障碍早筛语音数据”的产品，售价25万元，仅交易两笔。

图表 3 儿童构音障碍早筛语音数据截图



数据来源：蛋壳研究院

要解决这一问题，数据交易市场必须建立以政府为主导、市场化的数据要素交易机构和服务平台组成的体系。数据供应方进行有效的数据归集、加工；交易中心提供供需对接服务，实现数据升值、数据变现，监管机构保障市场监管和质控，营造良好的流通环境。

总的来说，我国数据要素流通市场仍处于发展的初级阶段，在数据交易需求持续高涨的趋势下，一方面需要国家主导完善数据要素服务相关制度，引导培育数据要素交易市场，另一方面也需培养更多供应商丰富数据供应体系，并加大数据交易所传播力度，最终实现在有效市场支撑下的数据供需匹配，让各类数据要素高效安全地流通起来。

# BIG DATA

## 第二章

基建扩容

千亿市场的根基

医疗大数据的价值挖掘是对医疗数据的收集、整理、再利用。换句话说，高效掘取医疗大数据价值，首先需要建立先进的信息技术系统和数据平台，帮助医疗机构有效存储、整合数据，其次才是构建、利用工具，对治理后的医疗大数据进行分析，进而为医疗决策提供科学依据、提升医疗效率。

相较于后一环节的医疗大数据应用，前端的治理显然更为繁琐。一个患者的全生命周期数据不仅包含产生于医疗机构的就诊、治疗、康复等临床数据，还包含与生俱来的基因组数据、可穿戴设备生成的个人健康数据。

以医疗机构产生的临床数据为例，患者在选择医疗机构时需考虑疾病严重程度、地理位置、优势学科等要素，一生之中的就诊不会局限于单家医院，因而其临床数据往往分散于多个医疗机构中；一家医院的临床数据分散在多个不同业务系统中，HIS、CIS、PACS、RIS之间未必能够实现互联互通。

有需求就有市场。为了治理过往沉积的数据，标准化后续生成的数据，不少医院开始投身于数据的数据中心的建立与数据平台的运营，两个模块如今已经孕育起一个十亿级的市场。

《全民健康信息化调查报告》曾对医院信息平台的基本功能统计。数据表明：已有超过一半医院进行了一定程度的数据中心部署。但在人工智能、物联网、语言模型等一众新式技术的推动之下，越来越多的数据中心面临升级、扩容、云化等需求，进而演化为新时代的医疗大数据市场。

图表 4 2021 年各类医疗机构大数据应用开展情况（单位：%）

序号	功能点名称	三级医院	二级医院	其他医疗机构
1	已开通	52.4%	59.9%	60.8%
2	数据交换	95.4%	85.6%	69.0%
3	数据质量管理	60.7%	49.7%	44.1%
4	数据查询	89.9%	93.5%	88.5%
5	标准字典库	78.2%	62.8%	48.0%
6	数据安全治理	55.3%	49.2%	44.1%
7	数据标准管理	68.3%	46.2%	49.1%
8	数据质量监控	47.7%	29.7%	25.6%

数据来源：《全民健康信息化调查报告》

## 2.1 新一代医院数据中心拉开千亿市场序幕

伴随信息技术和网络技术的跨越式发展,现代医院的运营特征表现为医疗业务的智能化与应用部署的敏捷化,使得医院业务产生的信息朝着复杂化、专业化、海量化的方向发展,并对各系统之间的互联互通提出更高层级的要求。

此趋势下,围绕网络带宽、服务器性能、交换机处理能力等设备运行特征建设数据中心机房的思路无法应对新型数据结构下涌现的数据安全保障、线上业务支撑、数据资产管理等需求,亟需引入新的 IT 架构来应对新的业务对计算资源、存储资源、网络资源的新要求。

因此,部分医院开始转变信息化建设思路,借助云技术打造新一代医院数据中心,综合利用各类数据服务临床、决策和科研过程,提高医院管理的科学化、规范化、精细化水平。

### 2.1.1 新一代数据中心的优势

传统大数据中心分为两类形式。一类是以业务支撑为主、整合电子病历的临床数据中心(CDR),其作用是支撑日常的医疗活动,收集与呈现医疗过程中的数据,绘制常规报表统计等。另一类以管理和科研为主的大数据中心,其作用是面向临床研究、医院管理与智能产品开发,满足科研、管理活动中的数据批量处理的挖掘与分析需求。

目前国内大部分全院级 CDR 完成了医院各业务数据的物理汇聚,但数据质量仍处于原始状态,对数据的深层架构与逻辑关系尚未进行梳理,针对现有 CDR 开展临床相关的数据分析挖掘仍具有极大困难。

此外,由于不同科研数据库一般采用自定义的数据模型,在建立多中心数据池、数据共享或数据合并时需要花费大量时间和资源进行数据映射和重新编码,一旦出错很容易导致计算机数据调用、分析过程和结果出现混乱。

要解决传统大数据中心存在的问题,新一代大数据中心应具备以下能力。

1. 满足医院创新业务需求。提供多种大数据应用开发工具并支撑大数据应用部署,例如利用 NLP 从海量电子病历数据中提取知识,辅助临床科研;利用深度学习从海量的医学影像中训练人工智能模型,辅助医生临床诊断。
2. 满足医院管理发展的需求。支撑人工智能应用为医院运营管理提供更深的洞察和更敏捷的反应;支撑实时流计算,能够应用大数据分析技术并将分析结果实时反馈到临床业务;支撑边缘计算与物联网技术实现智慧后勤。
3. 满足医院智慧应用配置需求。支持搭建安全、有弹性、可扩展的对外服务平台;支撑区块链等创新技术解决数据共享、流通、归集和安全问题。

4. 满足跨部门业务协同需求。支撑云网融合技术，能够在保证内外网间数据交换的安全性前提下以打通院内系统、外部系统及云上系统，以实现医院业务的连续性。
5. 满足数据治理需求。可提供的全局数据服务需要覆盖数据标准管理、基础数据管理、数据采集、数据汇聚、数据深度加工、数据资产管理、数据质量管理、数据安全管理等对数据的全生命周期治理服务。
6. 满足数据服务需求。支撑医院内部实现系统互联互通和数据对接共享的需要；医院提升海量数据资源质量的需要，数据驱动医院进行科学决策的需要，面对数据安全风险的需要。

图表 5 云上新一代医院数据中心主要特征



数据来源：《新一代医院数据中心建设指导》

### 2.1.2 新一代数据中心的架构

在《新一代医疗数据中心建设指导》一书中，新一代医院数据中心被定义为以私有云为主，多云结合为特征的医疗云数据中心，并将新一代医院数据中心的架构划分为医院基础设施服务层（IaaS）、医院应用支撑服务层（PaaS）、医院应用层（SaaS）。

图表 6 新一代医院数据中心架构



数据来源：《新一代医院数据中心建设指导》

在基础设施之上，医院应用支撑服务层提供平台应用服务，包括应用开发和测试的平台服务、通用应用服务、数据服务（DaaS）以及医疗行业特定的服务。

最上层的医院应用层提供各种类型的医院业务应用，主要分为三大类，即事务型业务、数据分析型业务以及跨院业务。这些业务应用通过平台即服务（Platform-as-a-Service, PaaS）层提供的 API 使用平台服务。此外，统一安全管理及统一运维管理将贯穿数据中心的各个层次，保证整个系统在安全环境稳定运行。

### 2.1.3 新一代数据中心的部署方式

目前，医疗健康行业对安全和隐私方面的高要求使得医院数据中心主要以私有云为主，这种方式不但可以保证医院完全控制其安全措施，也可以确保其数据中心能够满足相关法律法规。私有云部署也可以使医院明确知道自己所购买软硬件成本，使数据中心成本可预测。但是私有云部署的总体成本相对较高，也带来了医院管理的复杂性，而且数据中心的扩展性也受到医院本身条件的限制。

有些医院为简化管理，提高灵活性等因素考虑部分非核心业务也可能会选择托管公有云，由第三方服务商进行管理和运行，成本可控但其灵活性仍然有限。但出于对公有云可扩展性、高性能、低成本等优势的需求，医院越来越多地选择将其一些对外服务业务和 IT 能力部署在公有云上。

混合云技术的使用仍在探索中,该部署方式使得医院未来能够在私有云和公有云之间灵活地部署和迁移其工作负载,其典型的场景是将单个或多个私有云和单个或多个公有云结合为一体的混合云和多云环境。医院云数据中心也会通过互联网或专线与第三方平台对接,例如区域影像中心、全民健康信息平台 and 医保信息平台等。

## 2.2 影像补位, AI 企业成为基建主角

自深度学习广泛应用于计算机视觉后,围绕医学影像展开的人工智能开发催生了独立建设影像大数据中心的需求。因此,不少企业尝试在 PACS、RIS、PIS 等涉及医学影像的业务系统之上,以中台形式搭建一个能够综合管理应用全院所有影像的大数据平台,更标准、更便捷、更经济的方式对影像类数据进行汇总,进而推动相关科学研究的进行。

对比新一代数据中心,影像数据中心的需要同样宽泛。作为医疗数据大国,我国的影像数据占据了 80%—90%的份额,维持 30%的增速持续增长,但大量数据停留在纸质化的阶段,超过 80%的数据为非结构化数据。因此,围绕影像展开的医疗大数据中心建设能够有效触及此类需求,易联众、卡易、富士胶片等传统医疗 IT 企业及深睿医疗、汇医慧影等人工智能企业均围绕此业务深化信息化布局,一方面为医院搭载治理影像数据的平台,另一方面借助平台协同医院推出新的人工智能解决方案。

现有的医学影像大数据中心通常具备两类智能。其一为集成智能化影像应用。伴随人工智能技术的规模落地,一家医院常常会安装多个人工智能辅助诊断应用。为了方便医生在不同应用之间无感切换,医学影像大数据中心可以起到业务中台的作用,将众多人工智能软件集成,便于医生使用。同时,大数据中心还可以通过互联网实时更新新的人工智能应用,医生需要时可直接联网下载。

其二为助力影像数据资产化。由于影像数据量本身体量较大,医院很难精准评估影像数据集价值,可能在数年之后需要调阅时才发现文件发生损坏。在这一场景下,影像大数据可以助力影像数据治理,在数据生成后便及时打上标签并归类,并注明其可能存在的价值。

核心能力外,还有一些企业的大平台会打造一些特色功能。以深睿医疗新发布的 Deepwise MetAI 智慧影像&数据新平台为例,该平台以计算机视觉、NLP、深度学习等 AI 技术为基础,实现影像扫描后重建、打印、诊断、会诊、教学、科研的一站式影像科全周期智能管理。这个过程中,深睿医疗能够帮助医院生产高效精准的结构化数据,逐步积成影像科的优质数据资产。

同时,深睿医疗也在借助大数据平台打造以科研创新及应用为导向,实现数据到成果再到应用的创新闭环模式,加速科研成果转化,助力学科高水平发展。目前,深睿医疗已与四川大学华西医院、南京市中医院、浙江大学医学院附属儿童医院等国内多家医疗机构合作,依托



全院级科研数据中心，探索疾病智能化诊疗的新模式、产出高质量的成果、全面促进成果转化。

## 2.3 临床研发发力，科研药研加速数据库建设

临床科研数据库的建设需求基本可分为两大类：一是数据沉淀需求，通过积累数据资产，为尚未明确的研究课题和研究方向做探索准备，如医院对积累患者多组学数据的需求；申办方对积累临床研究数据资产的需求；积累医患管理数据的需求等。二是临床研究需求，如辅助已明确方向和课题的研究者发起的临床研究（IIT），以及药企发起的临床试验（IST）完成定向数据治理、质控、锁定和应用等相关工作。

人工智能、大数据等技术出现后，临床科研数据库的部署开始由“人力支撑”向“智能支撑”、从“强平台能力”和“强科研服务”的转变；专病库与临床研究数据采集系统（EDC）两类重要产品开始融合，逐步合并为一套系统。

临床科研数据库的价值随之深化。某肿瘤医院每年收治数千例宫颈癌、子宫内膜癌、卵巢癌等妇科恶性肿瘤新发病例，诊疗数据均存储在该院业务系统中，没有统一的标准规范和格式，且为非科研所需的结构化数据，非常不利于临床医生精准定位患者和统计分析数据。而在搭建了本地专病库后，患者群体定位能力和科研效率显著提升，数据处理成本降低，医院还获得更多科研机会。此外，该项目还反哺临床场景，帮助提升院内病历质量，增强规范诊疗能力。

图表 7 临床科研数据库的分类及特点

临床科研数据库的分类及特点						
数据库类型	研究区域	数据项及表单复杂度	数据标准化及质量控制要求	权限与追溯控制要求	接口需求及可实现程度	
临床试验数据库	单/多中心	低	高	高		中
病例注册数据库	多中心	中	中	中		低
专科专病数据库	单中心	高	低	低		高

数据来源：《临床科研数据库系统的现状与未来》

### 2.3.1 临床试验数据库

伴随数据科技的发展与医院对于临床试验数据库的日益重视，临床试验数据库的能力界限与发展路径都迎来的新的改变。

其一，融合院后随访数据。院后随访数据作为患者临床结局的反映，是多数临床研究中必须收集的数据。但传统的院后随访要么需要患者前往门诊主动随访，要么依靠临床研究人员电话随访。前者需要患者频繁往返医院，患者体验较差、后者存在随访工作量大、失访率高、数据质量受限等普遍性问题。

移动通信技术及应用广泛发展，临床研究人员如今可通过网页、微信、App 等多种形式提供患者院后随访服务以及收集患者自报告结局（PRO）。由于随访问卷、查体原始报告等信息可以由患者直接录入或上传，随访频率及随访质量均得到有效提高。

为了更好地融入院后随访数据，临床试验数据库应与新型的院后随访系统相集成，实现临床数据与随访数据的一体化整合，从而有效提升科研数据完整度。

其二，融合电子病历数据。临床科研数据采集需求会更多地体现在医疗业务信息系统中，两类系统之间的协同互补将是发展的趋势，但与医疗业务信息系统的协同，从医疗记录中回顾性提取科研数据，实践中始终存在着原始记录不全、结构化技术要求高的困难。为此，医院应建立“临床科研一体化”的电子病历系统，补全缺失的结构化数据。

但就国内医院目前的互联互通建设水平而言，绝大部分医院无法实现全结构化的电子病历。要解决这一问题，一方面可把结构化的科研数据表单以“插件”形式集成到临床电子病历系统中，兼顾临床记录的描述性和科研数据的结构化，另一方面可从专科化的检查方面推行结构化报告，逐步向全院进行延伸。

其三，完善数据加工功能。回顾性研究数据收集利用模式的改变在目前的科室专科专病数据库中，有相当一部分是为未来的回顾性研究而建设。一方面，在研究问题不确定的情况下，建设和维持科研数据库需要投入专门力量，长期持续难度大；另一方面，随着电子病历数据的完善和数据处理分析技术能力的提升，研究人员提出问题时，能够针对研究问题直接从原始数据中提取所需特征变量开展研究。

在专业化的数据服务能力支持下，直接提取科研数据有更高的效率，能最大程度地减轻临床科室工作量，这一趋势在医院的大数据中心工作实践中已经开始显现。未来，临床科研数据库系统的数据加工功能会进一步优化，入库的数据会更为精准、数目缩小，预先建库的模式将更多地转为有研究问题时的即时建库和数据加工，整个过程将进一步简化。

### 2.3.2 病例注册数据库

病例注册数据库主要用于支持疾病或治疗措施的观察性研究，通常为多中心数据库。研究设计者围绕研究主题和特定的研究问题，设计统一的病例数据采集表单，组织真实世界病例数据录入汇集，形成大规模的观察性病例数据库。现实中，一些队列研究数据库具有与病例注册数据库相似的特点，亦可归入此类。

病例注册数据库具有以下特点：通常为多中心数据库，对数据标准化定义要求较高；各个医疗机构的信息系统并不相同，病例注册系统与各医疗机构信息系统建立接口的可能性较低，数据采集主要依靠人工录入，因而病例登记表所涉及的变量数通常不宜设计过多；多中心采集场景下，对病例注册系统的数据录入和共享权限有特定要求。

### 2.3.3 专科专病数据库

专科专病数据库主要用于支持临床科室按专科或病种收集病例数据和开展观察性研究，其所收集的病例数据项较广。在实际工作中，科室建立专科或专病数据库可分为两种情况：一种是有研究目的但研究问题尚不明确，建库是为未来的科研问题积累数据，其所收集的病例数据项较广，病例表单构成较复杂；另一种是围绕特定科研问题的病例数据收集，数据项相对较少，病例表单相对简单。

随着医疗信息化的发展所提供的便利性增加，近年来此类数据库的建设需求越来越多。此类数据库与注册数据库的特点有类似之处，但由于属于单中心数据库且未来研究问题存在不确定性，所以多数数据库设计包含数据项较多，数据标准化定义要求相对宽松；对从医院电子病历数据资源库（CDR）获取数据的要求较高。

#### 2.3.3.1 专病数据库的构建趋势

传统临床科研数据收集因采取人工录入而存在数据格式、标准不规范以及利用率、成果转化率低等问题，无法满足日益增长的科研需求。因此，各分支学科均存在专病大数据建设需求，即通过建立专病数据库对医疗大数据进行整合与分析利用，提高专科疾病医疗服务质量及预测治疗效果，进而规避和降低医疗风险、抑制医疗成本等。

医学科技和信息技术的飞速发展，使得医院诊疗数据和病理及影像等数据，正在以惊人的速度增长。包括基因组、表观组、蛋白组和代谢组在内的组学技术迅猛发展，促使生物学领域快速进入“大数据”时代。依靠快速增长的生物学数据，重新审视疾病，对疾病进行细化分层，使得医学研究的重点更加精细和深入，逐渐集中于亚专业和专病，特别聚焦在高发病率、高死亡率、高疾病负担，严重影响人群健康的重大疾病上。

此趋势下，专病数据库的建设也需沿着新的建设方向发展。首先，建立目标专病数据模型及相关技术标准和规范专病大数据平台建设难点在于多源异构数据整合，需要建立通用数据标准，形成统一数据模型及相关技术标准和规范，以汇聚更多模态、来源更广的专病数据。

其次，应实现规范化数据汇集机制为对接医院业务系统临床信息进行底层数据集成，基于目标专病特点，利用标准化、规范化的数据元、术语库和同义词库，通过前置机及在线分析技术，根据机器学习智能推断数据结构、主外键关系、数据字典等元数据信息，采用自然语言处理、知识图谱和人工智能等技术对数据进行清洗、质量控制、结构化和归一化等处理，根据患者主索引及标准化数据模型建立肝硬化专病数据仓库，实现源数据自动化采集、数据解析与清洗、结构化与标准化预处理、基于逻辑与规则编写的指标抽取以及规则挖掘、知识查询、关联分析、图像处理、异常检测和预测分析。同时，需建立统一的规范化数据汇集机制，兼顾数据安全、处理性能和跨域传输能力，提供全量规则的医学临床真实数据。

最后，应建立高性能、高可靠性、高扩展性的存储架构目标专病大数据存储既包含 NoSQL 类型的数据库如 HBase、MongoDB 等，也包含 MySQL 等关系型数据库；存储架构引入 Redis、Memcached 等内存型数据库以提高实时计算速度；分布式计算框架则包含 Spark/Flink 流处理框架、Hadoop 等批处理框架、图计算引擎、数据挖掘引擎、TensorFlow 等人工智能处理引擎的多种分析框架，以满足不同分析场景需求。该存储系统具有分布式特点以应对数据规模增长；具有分层特点，即由高速和低速存储混合构成，高速存储保障在线实时或近实时分析，低速存储实现离线批处理等；同时具有完备的数据管理能力以满足数据冗余备份、同步、隔离等处理。

### 2.3.3.2 专病数据库的应用及效率评价

专科数据库支持医院各级医疗工作者和科研人员对目标专病的常规临床诊疗与医院管理所需的数据管理、查询、统计与可视化；可为临床与科学研究提供有力支持，可根据特定研究目标选定纳入标准、排除标准及输出指标来选择研究人群特征，以进一步在线特征描述、特征分析或下载相关数据开展更深入的数据挖掘和人工智能应用研究，例如大型队列、多模态数据融合的疾病预后、随访等临床科学研究；可支持基于标准应用程序接口的大数据分析挖掘及机器学习、深度学习等人工智能分析算法和模块的接入、嵌入，支撑未来数据驱动的目标专病临床与科学研究。

效率方面，过去通过编程方式实现 CRF 表单的设计与制作需要 1—3 个月的时间，而可视化、交互式的 CRF 表单配置可将时间缩短至 3 天；医生手工整理 1 份病历花费约 1 小时时间，而借助 NLP 处理，辅以手工填写补充，可将时间缩短至 20 分钟内。

### 2.3.3.3 尚需解决的问题

尽管科研专病数据库系统平台建设已为医院科研带来一定成效，但现有的信息化系统仍存在一定不足：

- 大量不同时期电子病历文件、检查报告、病理报告等非结构化的内容需要做结构化处理，但针对医学用语的 NLP 其数据理解能力仍有提升空间，对电子病历的语义分析有待加强。
- 现有的科研专病数据库系统中缺乏随访数据，后续还需加强互联互通建设，将专病数据库与随访系统对接，补齐科研数据短板，实现随访数据的共享利用。
- 数据库的维护和应用还不理想，建库容易维护难，课题结束后大部分的科研专病数据库处于无人问津的状态，造成新的资源浪费，需建立长期的数据采集和维护管理及奖励机制，鼓励临床医生和科研专病数据库研发建设的 IT 工程师，不断完善平台功能、定时补充更新数据，丰富科研专病数据库的内容。

## 2.4 数智化技术赋能，智慧后勤打开新需求

伴随医院的规模及能力日益增长，越来越多的数字化设备（医技设备、监护设备、智能楼宇和视频监控等）进入到医院。此趋势下，过去粗放式的后勤管理模式已经难以控制高涨的后勤成本。面对水、电、气、电梯、停车、物流、视频监控、高值耗材等要素组成的复杂的后勤体系，医院必须寻求创新技术实现更为精细化的管理。

要解决上述问题，首先需要医院进行有效的顶层设计，将繁多的设备统一至单个系统管理，并保证运营过程中的数据采集、数据治理、数据应用。人工智能支持下的 IT/OT 系统是一种理想的管理方式，该系统将信息技术（IT）与操作技术（OT）融合，并融入大数据技术进行包括物联网数据在内的异构数据存储以及快速数据访问和处理；借助边缘计算与物联网技术实现对智慧后勤业务需求的支撑。

不过，现阶段少有企业能够提供融合多重信息流的一体化系统。实际中，供应链、物流分属不同解决方案。峰禾科技、国医科技、海遇医疗、德荣医疗、万序健康、医贝、微萌、图特、爱惠、联众智慧等企业可提供 SPD 及对应硬件，赋能医院供应链管理；三维海容、瑞仕格等则可提供特定环境下的物流系统。

美的楼宇科技、达实智能等少数企业能够提供医院级的后勤系统。以美的楼宇科技旗下智慧医院 LIFE2.0 解决方案为例，该解决方案从交通流、信息流、体验流、能源流四个维度进行顶层规划，全面覆盖医院所面向的管理者、医护人员、病患和后勤人员等多维度人群不同空间的多元化需求。

图表 8 美的楼宇科技 LIFE 智慧医院框架体系



数据来源：美的楼宇科技

大数据工具支持下，医院各个场景产生的数据不再独立，且可实时传输至控制中心。在这一情况下，医院可以根据各个场景的运行情况调度及时调度资源，并根据历史时间数据对未来场景流量进行预测。

如在暖通工程这一场景，美的暖通设备匹配驱动算法软件有效联动，能够自动控制各等级手术室及病房的室内温度、空气湿度、空气洁净度、气流分布，提高患者就医环境的舒适度和医院用能效率，可将相关能耗降低约 30%。

总的来说，全局性的智慧后勤将为医院带来一场管理革命。但由于医院物流设计、暖通工程安装等环节对于医院的空间设计提出了很高要求，且很多医院在设计前并未考虑复杂工程的改造。因此，智慧后勤的市场规模会伴随医院院区的更迭不断提升，有望在未来持续发力。

## 2.5 大数据赋能的智慧医学教育

随着现代医学技术不断更新和创新，医学教育也必须不断创新，保证教育方式与教育内容的先进性，以便培养出具备先进医学技能和知识的医生。

智慧医学作为一种新的医学教育方法，采用了现代信息技术，结合大数据和人工智能技术，将医学知识技能和经验转化为数字化数据，已在教育和临床医学实践中丰富应用。相较于传统医学教育，智慧医学教育具备以下特征：

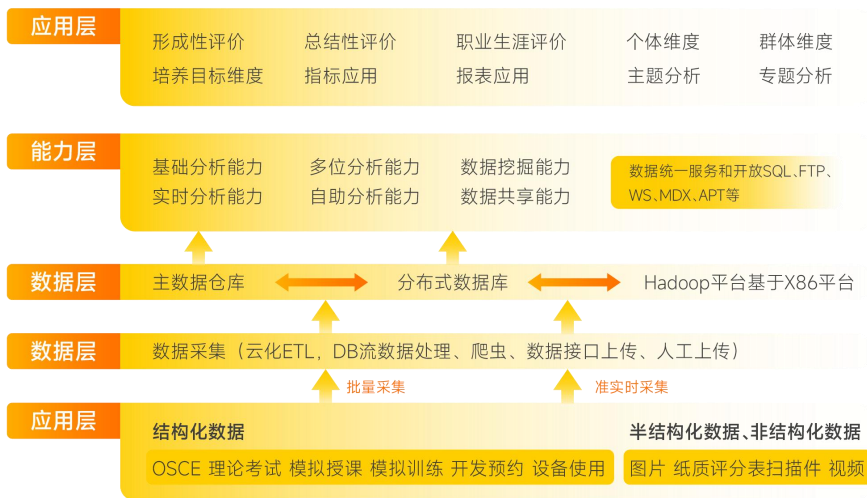
- 人性化教学。智慧医学在临床医学教育中可以根据学生的学习能力、兴趣和需求进行人性化教学。例如，使用虚拟仿真技术和实验室等工具和资源，对于不同的学生可以提供不同的个性化课程和培训模式。“虚拟患者”可以提供一系列的可控情境，帮助学生进行实践操作和审诊，同时又不会造成真实患者的危险和风险。
- 实时监测。智慧医学可以利用传感器和移动设备等技术实时监测学生的实践操作和诊断过程。这样可以在实践中及时发现学生的错误和欠缺，及时纠正和给出指导，引导学生不断改进和提高。
- 协同学习。智慧医学在临床医学教育中可以应用协同学习技术，使学生之间的互动和合作得到增强。这样有利于增加学生之间的交流，从而促进学术研究。我们可以结合成套的代表医学病例，对学生进行分组的探讨，进而学习各种临床技能，比如医学写作技巧、临床操作操场技，以及团队合作技巧等。
- 跨学科合作。智慧医学可促进不同专业之间的合作，进而为扩大学生的科学视野和增强交流平台出现，促进跨学科合作交流。临床医学教育可以将生物医学科学、信息科学和生物信息学等相关学科知识结合起来，增强学生的综合能力和创新意识。

目前，竹石数据、忆信捷等从事临床教学的信息化企业，以大数据、人工智能等技术应用于

临床教育医疗大数据平台的搭建中，柯林布瑞等企业也深入其中，尝试打造专注于教育服务的“EDR”。

以竹石数据为例，其大数据平台能支持学校和附属医院不同系统的多源数据对接，在架构上能满足对这些数据进行统一存储、处理、共享、分析、发掘；从业务上能支持学校对教育数据的最大化利用，实现数据的互通共享，科研、教学、管理等多个层次的使用，发掘教育教学规律，辅助教育决策，提升教育治理水平。

图表 9 竹石信息医教大数据平台整体架构及功能分层定位



数据来源：竹石信息

此外，该大数据平台能够对接和治理临床教学过程中海量的临床教学数据，并且能够使用流处理来应对学校系统中的一些实时数据（如设备和资源使用情况的监控数据），进而为学校提供功能丰富且安全可靠的教育大数据平台和相关的配套分析应用。

## 2.6 解答 5 个医院大数据中心建设中的常见问题

尽管大量医院已经意识到了医疗大数据的潜在价值，但在实际基础设施建设过程中，仍会存在一定问题。整合各位专家观点，蛋壳研究院将其梳理如下：

- 出于怎样的需求，医院需要建设新一代大数据中心？

临床数据中心，即临床数据存储库（clinical data repository, CDR），是一个通过标准信息表达和临床术语支撑的临床数据库，集成了来自医院不同临床信息系统的业务数据，实现多个业务域之间临床数据的存储。

过去，CDR 一直承担着日常业务的支撑功能，能够针对患者个体进行多方面的信息收集，实现医疗过程中的数据收集与呈现、常规报表统计等功能，具有实时性、长期性与稳定性。

相较之下，新一代大数据中心则是服务管理和科研过程中的数据分析处理，其作用是面向临床研究、医院管理与智能产品开发，满足数据批量处理的挖掘与分析需求。

目标对象不同外，两类数据库的差异还在于数据收集、工作模式、技术形态。

数据收集方面，大数据中心在进行分析工作时，不仅需要业务数据，还需要患者的随访数据、基因数据等。工作模式方面大数据中心的统计挖掘工作具备主观、大批量的特征，目的性较强，而 CDR 处理的工作具备日常、规律性等特征，义务性较强。技术形态方面，大数据中心需要分布式并行系统进行数据挖掘，而 CDR 会采用关系数据库，保障实时处理医疗数据。

总的来说。在以临床为中心的单一目的医院管理时代，医院建立 CDR 已经足够满足业务方面的需求。但随着科研需求的增加，新一代大数据中心将借助 CDR 的基础能力，在医院系统之中发挥更为重要的作用。

- 信息中心部门之外，医院是否需要另设大数据中心部门？

大数据中心和信息中心的职能与特点不同：传统的信息中心通常以信息系统建设和运维为主要职能，而大数据中心的职能是提供数据服务，尤其是为临床医学研究提供服务，这些服务是个性化的，需要特定技能的团队来提供。

大数据中心和信息中心的关系又非常紧密，大数据中心的建设紧密地依赖于医院的信息化建设，所有的数据都来源于信息系统，二者密不可分。另外在数据质量方面，也需要通过信息化建设来不断提升数据质量。

解放军总医院信息科高级工程师薛万国曾在 HIT 专家网上发表意见，他认为：理想的状态是信息中心提供原始数据，大数据中心分析数据、建立模型，所得成果再通过医院信息系统赋能临床。二者的紧密结合，可以在诸如静脉血栓栓塞症（VTE）风险预测、医疗质量控制、临床辅助决策等方面发挥重要作用。

他还表示：大数据中心和信息中心并非简单的分与合的问题，应该在医院内建立一个大数据体制，这个体制要体现出大数据中心是传统医院信息化内涵的扩展。至于大数据中心是否需要独立设置，这只是行政管理、分工管理上的形式问题。

- 专病数据库是不是数据研究的必由之路？

目前医院常见的专病数据库分为两种：一种是基础病例库，也即在建库时并未确定具体的研究问题，因此一般要求特征数越全越好，数据整理的工作量很大；一种是面向问题的病例库，也即针对特定科研课题的数据收集，对数据质量要求较高。

过往在建立专病数据库时，需要临床科室对所有病例数据进行人为加工，其优点是数据质量高、后期的数据处理量小，缺点是工作量巨大，需要建立长期的工作机制，临床科室往往很



难把这件事坚持下来。这也是大部分专病数据库效果不佳的主要原因。

薛万国认为：当医院建有大数据中心和相应的服务能力以后，可以将病例原始数据整合在数据资源池内，临床科研人员针对具体的研究问题，利用大数据中心提供的服务进行数据加工、特征抽取和数据分析。这种方式的优点是前期的工作量小、难度低，适合缺乏长期专门团队的临床科室，缺点是数据质量不及专病数据库。这种方式的好处在回顾性科研课题中表现较为明显。

- 数据安全问题能否解决？

在信息使用权限方面，医院大数据的信息安全领域并没有明确的法律规定，这就造成医疗大数据信息安全防火墙缺失，构成了信息安全隐患。医疗大数据的有关信息，也在一定程度上涉及个人隐私问题，因此必须对个人数据安全给予高度重视。

部分医院在推进医疗大数据建设的过程中，并没有全面认识到建立隐私保护、知识产权、数据信息安全等方面的医疗大数据法律与技术保障体系的重要性。因此，医院不仅需要加强对隐私保护立法的认识，为个人医疗信息的保密提供法律保障，还需要在居民健康管理过程中，明确信息使用权限。

医疗数据安全技术保障领域，部分医院的医疗大数据的安全和个人医疗数据的隐私保护同样需要加强建设。关于系统保护方面，部分医院在这方面的系统保护做得还不够。就目前来看，对于个人信息的管控还不是非常严格，“脱敏”“去标识化”处理不足，容易对个人隐私产生影响。因此，医院应建立安全信息通报制度，规范化风险隐患化解方案和应对工作措施。

- 是否需要独立招募大数据治理相关人才？

医疗大数据建设是一项具体的庞大工程，在这项工程建设中，许多专业性较强的事务需要大量的专业性人才来完成。当前，大部分医院严重缺少医疗信息人才，更缺少高素质的复合型信息技术人才，严重制约了医疗大数据建设。

由于医疗大数据建设中的技术工作对人员要求较高，一般技术人员难以胜任，导致医院内高水平的信息技术人员数量不足。因此，医院应招募特定人才执行大数据治理相关工作，保证建设的合理性与高效性。

# 第三章

应用创新

数字医疗建设进行时

*BIG DATA*

早期纷繁复杂的存储、清洗、分类，其根本目的是要构成标准化的数据集，服务于应用的打造及运营。因此，如何在构建医疗大数据集之后实现数据的高效应用，成为输出医疗大数据价值的关键所在。

《全民健康信息化调查报告》曾对现有的医疗数据的应用情况做出分析：现阶段医院各项大数据应用仍处于低位，三级医院应用数量占比不足 20%，二级医院应用数量占比不足 5%。其中，三级医院的运营、临床数据；二级医院的健康数据应用相对较多，其余数据仍需进一步发掘应用。

图表 10 各类医院大数据应用开展情况

应用分类	三级医院	二级医院	其他医疗机构
运营	20.6%	4.0%	2.1%
临床	21.9%	4.8%	4.4%
科研	10.5%	0.7%	0.6%
健康	7.6%	6.3%	6.3%
其他	1.5%	2.2%	1.9%

数据来源：《全民健康信息化调查报告》

缺失的应用比率正是科技医疗公司的机遇所在。本章将围绕大数据下的智慧应用建设，讨论科技公司在大数据时代面临的历史性机遇。

### 3.1 临床大数据应用

作为数据资产化中最具价值潜力的一类数据，临床信息化建设的进步不断推动基于大规模医疗数据的临床真实世界研究，以电子病历为数据源建立临床研究数据库的需求愈发繁多，临床科研数据库系统的功能需求及支撑技术也在不断发展演化。因此，在传统的病例数据收集和利用模式发生变化的背景下，总结分析临床科研数据库系统的发展对于完善临床科研数据库建设、提升临床科研支撑水平、打造基于临床大数据的相关应用具有重要意义。

#### 3.1.1 临床辅助决策系统

根据《临床决策支持系统的构建与应用》的定义，临床决策支持系统（Clinical Decision

Support System, 简称 CDSS) 是指运用系统的临床知识和患者基本信息及病情信息, 加强医疗相关的决策和行动, 提高医疗质量和医疗服务水平的计算机应用系统。简而言之, CDSS 的出现是为了帮助医院通过“数据治理”触达“临床管理”。

CDSS 的结构通常由知识库、推理机和人机交流三个部分组成。最早的 CDSS 又称专家系统, 虽也能对已知信息和数据的分析与解释, 确定它们的含义, 但由于数据库包含数据有限、逻辑推理规则较为简单, 这类一方面严重依赖专家经验, 且无自主学习能力, 难以应对复杂问题。

人工智能与大数据相关技术及相关企业的出现重新定义了 CDSS 的能力, 推进了专科 CDSS 的发展。数据层面, 卫和医学等权威知识库提供方以实时更新、自动化处理、可解释内核的动态数据库取代了传统静态数据库, 为知识图谱的构建提供的权威全面的医学知识支撑。

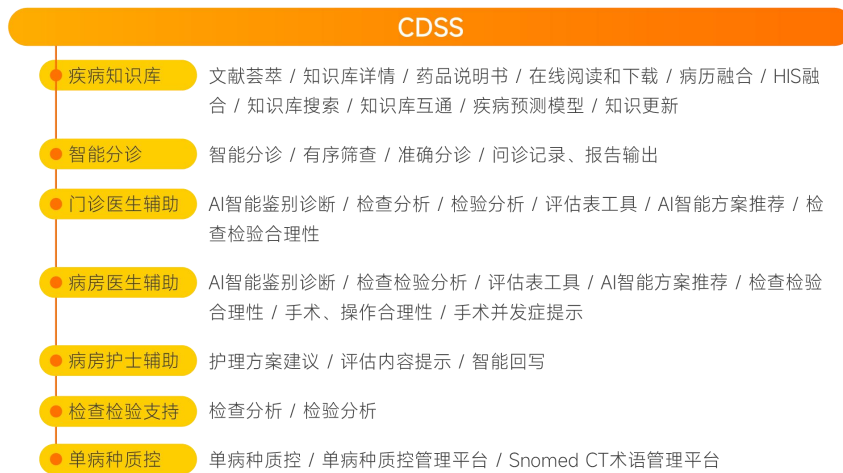
图表 11 卫和数字生命体征医学模型



数据来源: 卫和医学

算法层面, NLP 等智能技术的崛起, 使得 CDSS 能够清晰“理解”需求, 为临床医生提供诊断、治疗、护理、手术、合理用药等方面的决策支持; 为具体疾病提供建议、提醒、报警、计算、预测方面的决策支持。

图表 12 医院采购 CDSS 产品涉及的主要系统



数据来源：蛋壳研究院

现阶段下，CDSS 需要突破的难点大致分为两点。其一，全科 CDSS 已经很好地将权威的知识库融入了模型中，但绝大多数的 NLP 仍需提升语义理解能力；其二，专科 CDSS 已经在 VTE、房颤等医疗场景中获取一致认可，但受制于数据获取、场景认知等因素，创新产品的开发速度仍有待提升。

### 3.1.2 人工智能辅助诊断系统

人工智能的能力界限取决于算法、算力、知识、数据四大要素。对于初创公司而言，算法可以借助开源模型修改，算力可以购置 GPU 获，知识可以寻求权威知识库合作。唯独数据，既需要企业同医院取得合作，协助治理数据，建立高质量数据集，又需要导入模型，逐步调试参数，取得合适的训练结果。

与专科 CDSS 面临的问题类似，创新人工智能辅助诊断系统的开发同样受制于病种数据的限制，因而医疗大数据的快速发展将帮助 AI 公司以更低成本获取更高质量、更大规模的数据，突破现有应用场景的局限性，进而颠覆 AI 赛道的发展。

伴随影像数据标准化的推进，人工智能辅助诊断系统已经突破了眼底、肺结节等传统应用场景，实现了多模态、多病种、全流程的覆盖，甚至深入治疗领域，辅助手术导航。

以沛心科技的“智心”CardioVerse 为例，该 AI 是一种基于 CT 影像数据的瓣膜评估及手术规划系统，可帮助术者进行手术风险评估，指导手术策略，保障 TAVR 手术安全性，提升手术成功率，助力 TAVR 术式下沉。

推想医疗的“龙点睛®穿刺手术机器人”则瞄准了经皮穿刺，利用人工智能技术，融合磁、机、电、控、软、算等多学科，实现全自动组织病灶识别、手术自动路径规划和穿刺引导及消融规划评估，帮助医生准确定位病灶、提高穿刺准确度、减少患者手术创伤、缩短手术时间、降低术中和术后并发症的发生、减少术中 CT 扫描次数和辐射，从而辅助医生更准确、更快

速的完成经皮穿刺下的早期癌症消融的手术操作。

### 3.1.3 真实世界研究

真实世界研究是指针对预设的临床问题,在真实世界环境下收集与研究对象健康和疾病有关的数据(真实世界数据)或基于这些数据衍生的汇总数据,通过分析,获得药物或相关器械使用情况及潜在获益—风险的临床证据(真实世界证据)的研究过程。真实世界研究可以是观察性研究,也可以是干预性研究。

与传统随机对照试验(randomized controlled trial, RCT)不同,真实世界研究的数据来自真实临床场景,证据外推性好,可用数据量大,研究易于开展,成本相对较低,可帮助研究者发现临床实际情况与理想 RCT 研究之间的差距。

有效的真实世界研究一定程度依赖于研究者对于医疗大数据的分析管理能力。在构建数据集时,研究者应执行可行性评估、数据质量控制、方案分析、方案评价、输出与反馈、安全与合规六个环节,过程中需要完成数据评估、数据获取、数据存储、数据清洗、模型建立等工作,非常依赖大数据平台的处理能力。

## 3.2 运营大数据应用

将大数据分析技术应用到医院运营管理当中,能够在海量的大数据中挖掘出最具价值的信息,保证医疗企业中的各项管理工作得以顺利展开。在相关医疗 IT 企业的支撑下,部分医院经营数据已得到了充分分析应用,管理人员结合大数据分析平台提供的各项数据,全面了解医院各科室经营管理情况,对原有的经营管理制度及经营状态进行实时调整。

### 3.2.1 为 DIP 提供技术支撑

基于大数据的病种组合(DIP)是利用大数据优势建立起的完整管理体系。该体系应用“随机”与“均值”的经济学原理,借助真实海量的病案数据,发现疾病与治疗之间的内在规律与关联关系,提取数据特征进行组合,并将区域内每一种疾病与治疗资源消耗的均值与全样本资源消耗的均值进行比对,形成 DIP 分值。目前, DIP 主要适用于住院医疗费用结算,精神病、康复类及护理类等住院时间较长的病例不宜纳入,其适应性及扩展性可探索应用于门诊付费标准的建立。

DIP 与 DRG 的逻辑大致相同,均以将资源消耗相似的病例进行聚类作为理论基础,形成若干病种组。但 DRG 分组按照 MDC-ADRG-DRGs 的三层逻辑,一般只包含 600-800 组,而 DIP 以一级至三级目录递进的方式完成建立支付病种表,二级目录约 3000 组,三级目录共计 16000 组。此外, DRG 从医学理论出发,立足于数据分析,每一组都严格遵照从解剖系统,到疾病治疗方式,再到病案个体特征的分层逻辑;而 DIP 则是基于客观的大数据事实,

即病种“存在便为合理”。

目前，DIP 的落地方式通常分成两步走：第一步仅基于主目录，即主诊断+主操作+病例数进行落地，不考虑次要诊断，肿瘤转移，放化疗及离院方式等因素，故对数据质量要求较低。第二步则需加入病种第三层辅助目录进行考量，即肿瘤程度及病情程度，对次要诊断填写要求较高。

因此，信息化水平较差的地区可以率先开展第一步，待基础数据质量得以提升后，再对病种进行细分更新。这种模式下，支付改革的进程就不会受到数据质量的制约，更容易在数据基础建设较差的医院进行推广。信息化水平较高的地区可直接进行第二步，快速完成 DIP 部署。

### 3.2.2 以 DIP 辅助医院运营决策

DIP 对于医院运营决策的促进作用主要分为以下三点。

一、DIP 促进医疗机构加强费用控制，提高经济运营效率。DIP 结算对样本医院的经营管理提出了较高要求。样本医院围绕病种分值与单价、时间与费用消耗指数等关键指标，建立院内运行评价体系，提质控费，促进科室提高经济运行效率。样本城市医疗保障部门要求，医疗机构每年人均费用增长不高于当地 GDP 增幅，同时将费用消耗指数及时间消耗指数两项资源消耗指标纳入考评范围该院多部门协同，紧密围绕 DIP 病种，进行成本效益分析，努力降低病种费用。

《DIP 改革赋能三级医疗机构发展内涵的实践与应用研究》以子宫多发性平滑肌瘤行腹腔镜子宫病损切除术为例分析了 DIP 的价值。该研究将某医院的子宫多发性平滑肌瘤行腹腔镜子宫病损切除术病历分为五组，应用 DIP 后计算费用结果。数据显示：不同诊疗组之间盈亏状况具有明显差异 ( $P < 0.05$ )。诊疗组 C 盈利 1245.53 元，诊疗组 E 亏损 1263.14 元，诊疗组 C 与诊疗组 E 的药品折合分值分别为 35.93、44.47，耗材折合分值分别为 53.47、60.46。亏损诊疗组药品、耗材分值均高于盈利诊疗组，不同诊疗组间药品、耗材分值具有明显差异 ( $P < 0.05$ ) 该病种成本结构可以进一步控制、优化。

图表 13 不同诊疗组子宫多发性平滑肌瘤行腹腔镜子宫病损切除术费用比较 ( $P < 0.05$ )

类别	诊疗组	N	均值	F	P
药品分值	诊疗组A	57	41.42	1105.8%	0.000
	诊疗组B	50	37.08		
	诊疗组C	17	35.93		
	诊疗组D	27	33.18		
	诊疗组E	33	44.47		
耗材分值	诊疗组A	57	56.49	327.9%	0.013
	诊疗组B	50	51.84		
	诊疗组C	17	53.47		
	诊疗组D	27	52.15		
	诊疗组E	33	60.46		
医保亏损金额	诊疗组A	57	-403.61	1092.2%	0.000
	诊疗组B	50	-405.13		
	诊疗组C	17	1245.53		
	诊疗组D	27	1940.70		
	诊疗组E	33	-1263.14		

数据来源：《DIP 改革赋能三级医疗机构发展内涵的实践与应用研究》

## 二、通过入组规则及辅助目录促进医疗机构提高病案编码与质控水平

DIP 严格按照诊断编码前四位及手术、操作编码组合入组，同时考虑患者具有多项并发症、高龄与疾病严重程度等因素，建立辅助目录内容包含疾病严重程度、肿瘤 严重程度分型、次要诊断病种\_年龄特征病种、CCI 指数等多项内容。上述规则促使医疗机构加强病案首页质控管理，准确填写各项诊断与操作，不漏填、少填，争取医保费用得到合理结算

## 三、促进医疗机构提高 CMI 值，提升核心竞争力

DIP 支付体系鼓励了三级医疗机构严格控制基层病种的收治，提高疑难危重患者收治比例。对于费用极高患者，医保实行特例单议。医保多种措施促进医院调整病种结构，不断提升核心竞争力。DIP 付费实施以来，样本医院按院一科一组能力评价结构，确定“学科、技术、质量、费用、效益”五个维度评价的《基于病种管理的临床发展能力评价体系》，共 16 个二级指标、51 个三级指标，全面评价临床科室发展能力，确定优势病种，推动价值医疗，构建医院学科发展新格局。

### 3.2.3 DRG/DIP 助力医院国考

在 DRG/DIP 发展过程中，大量医院为满足政策要求将 DRG/DIP 体系的能力范畴局限于第一维度。而医保数据贯穿不同医疗业务单元，其应用场景串联全院管理链条，它不仅仅可充当



医保控费的工具，亦有能力激活全院数据智能，成为医院精细化高质量“转舵”的抓手。

因此，在第二维度中，医院需要与资源消耗管理、精细运营管理、临床路径管理等路径深度结合，帮助医院从全过程规范管理，到一体化精细运营，直至以资源消耗路径优化临床路径，实现临床诊疗“质效”融合提升。

以国新健康的“DRG/DIP+”解决方案为例，该方案对于首先是数智赋能医院全过程规范管理。通过将病案质控、诊间监管、医保结算清单质控融入 DRG/DIP/APG 医院智能管理系统，实现医生填写病案时即开始审方、监管、质控，并利用预测性分析技术，根据即时诊断和手术预测分组展示不同分组方案，帮助医院在 DRG/DIP/APG 预分组、支付标准和结算差异等方面前置干预，过程预警。在动态监测、靶向定位可能存在的监管风险后，适度、规范管控，一站式结算—质控—上传—归档—反馈—申诉，保障数据上传及时、准确、全面，提升医院一体化管理能力。

其次是“DRG/DIP+”精细化资源消耗管理驱动精益运营。面对医院医保医疗服务与运营管理的痛点及实际需求，基于丰富的病案质控规则，国新健康“DRG/DIP+院内绩效分配”针对医保支付和卫健考核形成一体化解决方案，深入各业务场景持续挖掘分析，聚焦问题、落实精准，有效帮助医院将绩效考核指标转化为业务运营管理。在此基础上，其“DRG/DIP+医院全成本核算和预算管理”，从院级成本、科室成本、项目成本向病组成本和病种成本深化，促进临床诊疗规范，并挖掘医院优势病种，助力形成特色学科。

同时，国新健康依据成本核算数据做出成本预算，将预算管控点前移至业务端，在符合临床服务与科研教学业务发展要求、人财物联动资源配置需求的前提下，预警医生端药耗使用、监测运营端设备资源配置，并融入绩效考核中，从而激励医务人员主观能动性，助力医院逐步从“收入增长”转变为“降低成本”，从财务视角转向价值链视角，全面、精准地实现以成本核算为基础的医院精益运营。

最后是优化临床路径实现质效融合。DRG/DIP/APG 在院内形成的是资源消耗路径，以资源消耗路径优化临床路径，实则通过对 DRG/DIP/APG 的指标分解，从服务能力、服务效率、质量安全三个方面，建立一套诊疗全过程管理工具，重塑“质量—效益”管理新模式。在 DRG/DIP/APG 的政策触发下，医院将以提升药物合理使用、提高检查检验合理使用、加强平均住院日管理、加强超支病组管理等为关键抓手，打通以临床业务为源头，通过合理诊疗、因病施治，改善资源配置，提高服务质量，激励协同合作，进而反哺临床的 PDSA 医院一体化精细运营管控路径，实现兼顾医疗质量与卫生经济双提升的“质效”融合。

总的来说，国新健康的方案立足于医保支付本位、深入医疗业务及管理全过程、扩展“DRG/DIP+”一体化运营服务的行业布局，有助力医疗机构逐步形成数据采集标准、使用规范统一，以病种和临床路径为单元、医保支付为基准、全成本核算为结果、价值医疗为导向

的全面精益运营管理体系，进而实现三级医院“国考”排名的有效提升。

### 3.3 院外其他大数据与应用

相较于院内大数据，院外大数据归属于数据生产方，产权清晰，能够更加针对性地构建平台，开发应用。因此，不少 IT 企业针对药房、保险等机构设计大数据应用，提升经营管理效率，缩减销售运营成本。

#### 3.3.1 药店智慧管理

新冠病毒的肆虐为社会带来了许多消极影响，但也推动了不少互联网技术的进步。在药店领域，消费者对于智慧药店的认知大幅度进化，进而也倒逼药店经营者理念发生转化，智慧门店无疑成为新零售的前台和主战场。

该形势下，药店同时面临挑战与机遇，一方面需要利用人工智能、大数据等技术建立智能化平台，通过分析各类数据提升连锁药店管理能力，另一方面可利用现有互联网资源帮助连锁药店探索新的盈利途径。

部分企业协同智慧药房为患者提供个性化、专业化的健康管理服务。通过患者的身体数据和健康状况等信息，智慧药房可以为患者提供更加全面、细致、贴心的健康服务，例如为患者推荐合适的药品、健康饮食和运动计划等，扩展药房业务范畴。

也有企业借助大数据技术及模型创新着力极具潜力的创新药院外市场，打开药房销售新增长点。譬如，上药云健康通过树立“益药”全国一体化品牌，以构建专业药事服务能力为基石，全渠道打造全国领先的一体化专业药房体系。2021 年上药云健康收购融合百济新特药药房、康德乐大药房；同时整合上药院边药房，纳入“益药”网络，DTP 业务进一步整合，逐步形成以“益药”系列为核心的全国新特药销售与服务品牌，已覆盖全国 25 省 66 市。

作为中国最早开展新特罕专科病种服务的专业药房，上药云健康“益药·药房”经营 6 大病种科组，130+个重大疾病和慢病病种。拥有业界认可的培训团队、药师团队及患教团队，以更好地为患者提供标准化、专业化的药学服务，确保用药的安全性和有效性。“以患者为中心”，通过药学服务准备（内容建设、药师培训）、标准化患者服务过程、专业化药物治疗管理（线上科普、患者治疗跟踪、信息完善和数据分析、患教活动）实现全程化的药学专业服务路径，提升患者的治疗信心和用药依从性。

#### 3.3.2 DTP 助力 RWS 研究

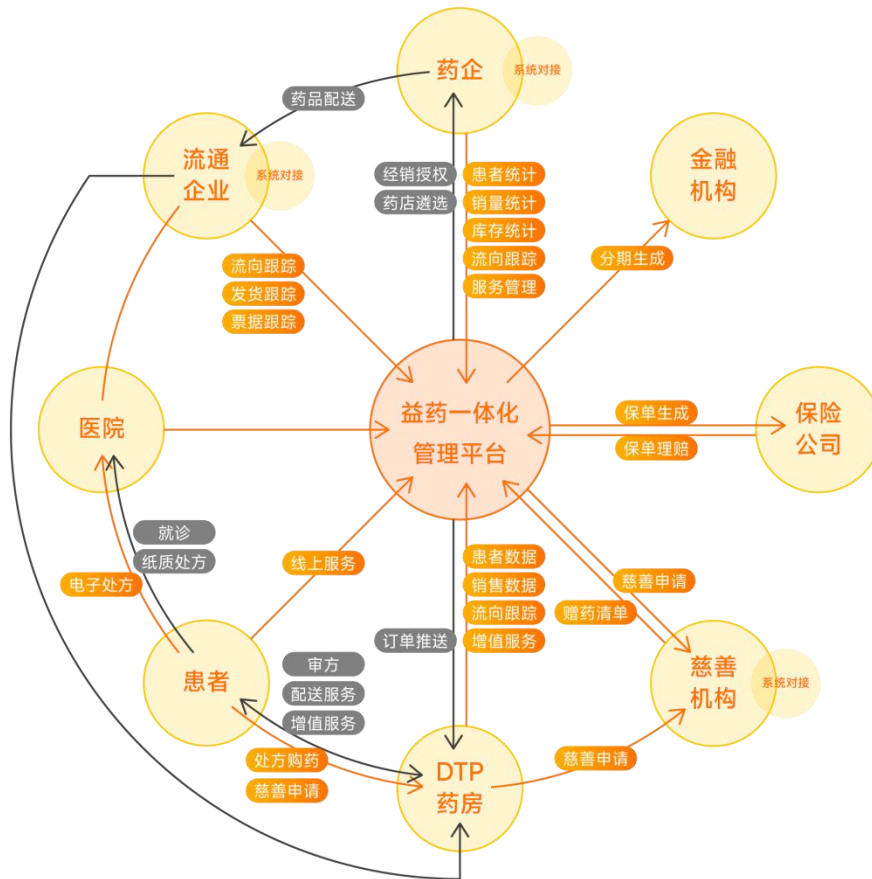
与传统零售药店模式不同，DTP 药房销售主要是单价较高、存储条件较为严格的新特药，主要包括肿瘤药、罕见病用药等，且构建了一个连接医、患、药、保的综合服务平台，以病人为中心提供专业化、个性化、全周期的健康管理服务。

同时，作为院外重要的患者管理与服务场所，DTP 药房对肿瘤患者的全程管理具有重要的作用和价值。一项在美国西维斯特药房开展的回顾性研究显示，在药房药师和患者间建立双向的信息沟通模式可有效提高慢性粒细胞白血病患者治疗依从性。相比之下，国内与 DTP 相关的 RWS 较少，但伴随上药云健康、思派健康、零氦科技等企业的介入，相关探索有所增加。

2021 年思派健康旗下思派大药房联合北京大学医学部展开国内首份《DTP 药房患者管理服务效果的真实世界研究》，该《研究》采用回顾性队列研究设计方法，利用全国 79 家思派大药房的销售及随访数据，对 2019-2020 年间购买 PD-1/L1 类药物患者的购药和随访数据进行回顾性分析。研究结果显示，与随访接通率小于 30% 的人群比较，随访接通率大于 70% 的患者治疗中断及随访中断的风险分别下降 21% 和 42%；治疗中出现不良事件的患者，中断治疗及随访的风险分别下降了 24% 和 31%。

上药云健康同样在专业化患者服务和提升患者依从性方面进行了广泛研究。该企业打造了 DTP 一体化平台，拥有 60 余个品规，链接 600+ 家自营以及合作 DTP 药房。通过完整的数据链条和全景图，为药企提供患者疗效分析报告、患者黏性报告、用药人群画像、渠道分析报告、销量分析报告等全生命周期患者旅程服务。未来，基于与药企长期在临床领域开展对照药、器械、患者招募等方面的合作，业务触角将进一步向全周期数字化赋能方面延伸。基于药房网络和线上平台，汇集完善的患者数据池，通过专业的药事服务，增加与患者的触点，提升患者洞见的深度和广度，优化患者服务能力。

图表 14 上药云健康益药 DTP 一体化平台



数据来源：上药云健康

### 3.3.3 基于大数据的健康险定价及风控

历经 30 余年发展，我国健康险行业已经取得长足进展，2019 年健康险原保险保费收入达到 7066 亿元，近十年的年复合增长率超过 28%。

健康险，尤其是医疗险，非常考验企业精细化运营，过程中的风险定价需要大量的数据支撑。但由于缺乏医疗数据对接和共享机制，精算数据库无法得到扩充，因此健康险的产品创新受到抑制，同质化现象愈发严重。

以惠民保产品举例，本应“一城一策”、根据不同地区居民健康状况来定价，但在市场白热化竞争的驱使下，为了尽快抢占市场，不仅条款相似度非常高，定价缺乏足够的医保数据支撑。

同时，在基本医疗保险形成广覆盖、保基础供给的同时，商业健康险的发展相对滞后，没有完全发挥其市场灵活性，在产业链中的地位也相对弱势，尚未能借助其支付者的优势串联起医疗健康产业链。

在经营管理上，许多险企也很难脱离“泛寿险化”经营的惯性，产品结构单一、渠道开拓使用成本高、专业化水平低、风控手段滞后、盈利难度高等问题一直未能得到很好地解决。

从长期角度看，商业健康保险需重新定位其医疗健康产业链支付角色这一优势，通过串联用户、医、药、健康管理等多方利益体，构建医疗健康生态价值链，形成事前预防、事中诊疗和事后管理的闭环健康服务；同时通过与基本医保的差异化定位、充分依托大数据技术优化行业经营，寻找新的利润空间，改变被动支付、依靠传统三差的单一盈利模式，探索新的商业发展模式。在广度上，构建覆盖全渠道、全人群、全生命周期的健康险产品供给体系。在深度上，依托社保目录到建立供需联动、风险可控的动态商保目录。

在生态上，变基础采购为融合共生，平衡和重构保障边界带来的风险成本。在风控上，串联医疗资源与商保供给，创新支付模式，建立风险共担机制。实现上述转型发展，一个重要的依托就是健康医疗大数据。在以前信息化和数字化程度还没有达到极大丰富、互联互通的时候，主要是以人工方式来搬运信息；随若医改深化和医保信息化水平不断提升，社商数据联通成为可能，这就让商业健康险有机会走出现有的经验逻辑，真正地从客户保障需求出发，建立面向不同用户的、精准定制化的全生命周期保险保障供给，基于数据突破现有的以健康险体为主的经营瓶颈，从“保健康人”走向“保人健康”，构建新的健康管理和医疗支付的生态供给体系。

此外，健康保险业务对客户个人信息的安全和保密性要求非常高，现在越来越多的健康数据可供利用，如智能手表、健康监测器等，这些设备可以收集到客户的身体健康状况，如心率、血压等。如果可以将健康数据与个人风险结合，保险公司可以更准确地评估客户的风险水平，并为客户提供个性化的健康保险服务。

### 3.3.4 公卫大数据的应用

科学高效的突发公共卫生事件防控措施是应对突发公共卫生事件的关键。在突发公共卫生事件爆发后，能够基于联邦学习，关联医疗、交通、通讯、教育等多源数据，可视化展示疫情发展相关数据和知识，并支持大规模核酸检测、密接和次密接排查、社区隔离管控、患者治疗、应急物资调配、疫苗接种等。如在患者治疗过程中，通过联邦学习开展跨地区、跨机构协作共享，涉及病人隐私的电子病历、电子健康档案、病症、病理报告、检查检验结果等数据，各级医疗机构可实现共享，通过患者症状、检查检验结果、诊断、治疗等数据，运用决策树算法，发现大量病例中蕴含的规律，建立疾病诊断和治疗模型，辅助医生疾病诊疗。

在应急物资调配中，需要涵盖各级卫生健康行政部门、政府、疾病预防控制机构、医疗机构、街道社区等庞杂的机构体系，如何动态掌握应急物资储备情况、缺口种类和数量，最大限度发挥应急物资使用效能，是保障突发公共卫生事件防控的关键。基于联邦学习，能够在保障数据安全共享的基础上，实现应急物资状况的精准监测，并能够通过人口分布、人口结构、人口流动情况、感染人数、死亡人数等指标对各类应急物资使用情况进行预测。

传染病监控对传染病报告卡的数据质量要求较高，包括传染病报告的及时性、传染病报告数据的完整性和准确性都是重要指标。这本是一件好事，但客观上也使得医生在实际传染病报卡过程中会遇到很多问题。

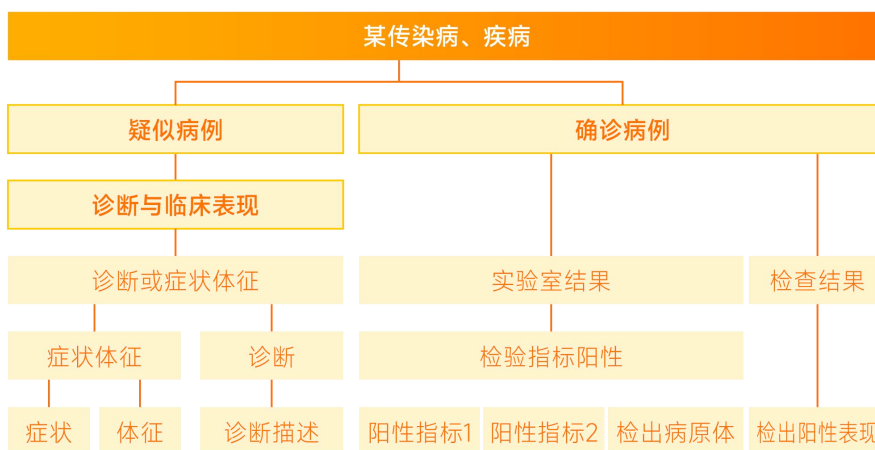
首先是及时性。现行规定甲类传染病必须在 2 小时内，乙类传染病及丙类传染病须 24 小时内上报。然而，医生填写传染病报告卡信息需要花费较多时间，遇到就诊病人多时，就容易出现疏漏造成没有及时上报，这样就会导致传染病的迟报和漏报情况的出现。

其次则是传染病报告数据的完整性和准确性。目前多数医院挂号系统中并没有完整的地址及电话等信息，在填报时无法满足传染病报卡对病人信息完整性的要求。医生只能现场询问并录入，过程耗费 3~5 分钟时间甚至更长。

最后，医生对传染病疫情的认知也影响到传染病的监测。在疾控直报专网的传染病报卡中有“不明原因肺炎”的选项，实际上可以对应早期未明确的新冠肺炎。现实情况则是医生基本上都没有选择此项选择报告，因为上报就意味着后续需要完成一系列登记填表及调查工作，对于他们是额外的负担。或者，即使知晓有此选项，但缺乏明确的诊断也不敢轻易上报或无法通过现有系统的规则进行上报。

对于上述问题，大数据、人工智能等前沿技术组成的融合方案或能一定程度解决上述问题。目前，国内已有和宇健康、双数科技等企业的系统入疾控中心。以双数中心的传染病疾病智能监测方案为例，其利用大数据和人工智能的双数传染病监测方案可以将杜绝迟报漏报的有效率做到 95%以上；将医生的平均传染病报卡所耗时间从以往的 5—8 分钟大幅降低到 40 秒以内；疾控直报专网上报所耗时间也从 2~3 分钟缩短至几秒钟，有效应对上述问题。

图表 15 传染病疾病智能监测方案运行逻辑



数据来源:蛋壳研究院

# BIG DATA

## 第四章

保障数据安全

构架医疗大数据价值流通的基础

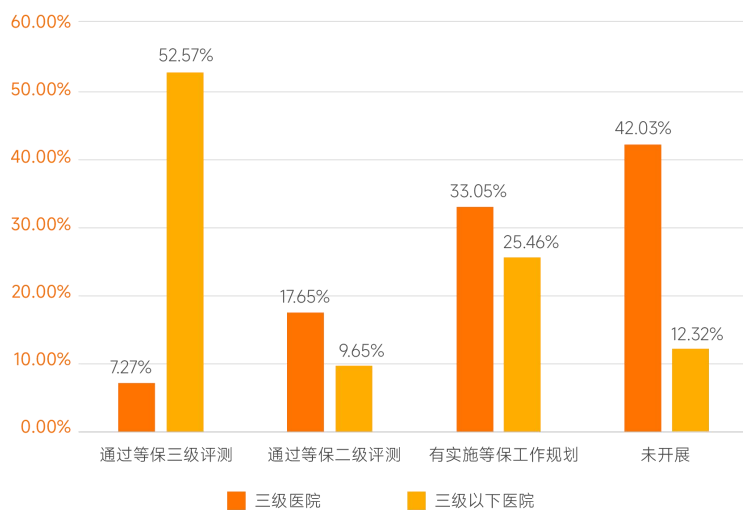
与劳动力、资本等生产要素相似，医疗数据的价值体现于应用与流通过程。但无论是在数据平台间应用流转，还是未来他通过交易所变更归属主体，只要存在流动，必然会数据泄露、数据滥用等问题。因此，只有保障医疗数据流动的安全，才能保障医疗大数据产业的良性发展。

## 4.1 医疗大数据安全现状

为推动医院加速信息安全基础设施，我国从 2011 年便开始陆续推出相关政策，一方面借助《卫生行业信息安全等级保护工作的指导意见》（卫办发〔2011〕85 号）、《关于印发医疗质量安全核心制度要点的通知》（国卫医发〔2018〕8 号）、《关于印发全国医院信息化建设标准与规范（试行）的通知》等一系列文件，以等级保护建设为中心推动医疗机构网络安全建设；另一方面在《电子病历系统功能应用水平分级评价方法及标准（修订征求意见稿）》《国家医疗健康信息医院信息互联互通标准化成熟度测评方案（2017 年版）》《国家医疗健康信息区域卫生信息互联互通标准化成熟度测评方案（2017 年版）》等多项评级要求中提出目标，将宏观目标拆分为微观细则督促医院执行。

作为这一无形核心资产的持有方，医院对于信息安全的关注度也在持续提升。CHIMA 调研数据显示：2018-2019 年统计的三级医院通过等级保护三级测评的比例为 52.57%，三级以下医院通过等级保护测评（包括二级和三级）的比例为 24.92%。多数医疗机构，尤其是三级以下医院网络安全等级保护建设仍处于较低水平，整体未开展比例为 26.87%。

图表 16 2018-2019 年度中国医院信息化状况调查医院等级保护工作情况不同等级医院对比



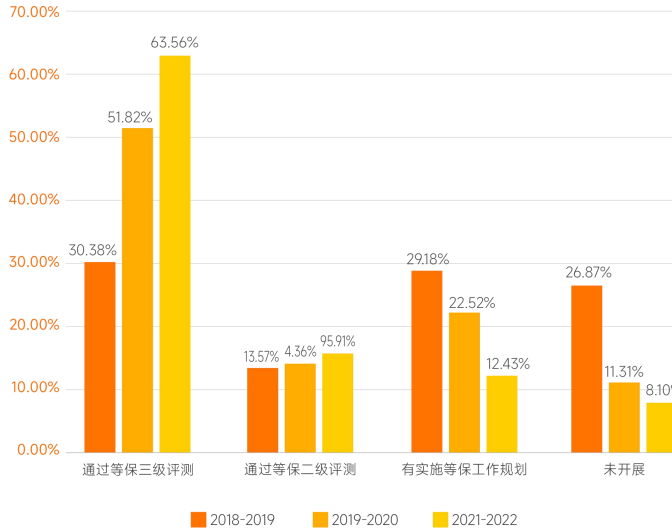
数据来源：CHIMA

政策支撑后，2022 年通过等级保护三级测评医院已经到达 63.56%，较 2019 年翻了一倍有



余，未开展三级以下医院仅有 24.92%通过等级保护测评（包括二级和三级），整体比例也从 26.87%缩短至 8.1%。

图表 17 医院开展等级保护工作状况不同年度对比



数据来源：CHIMA《2021-2022 年度中国医院信息化状况调查报告》

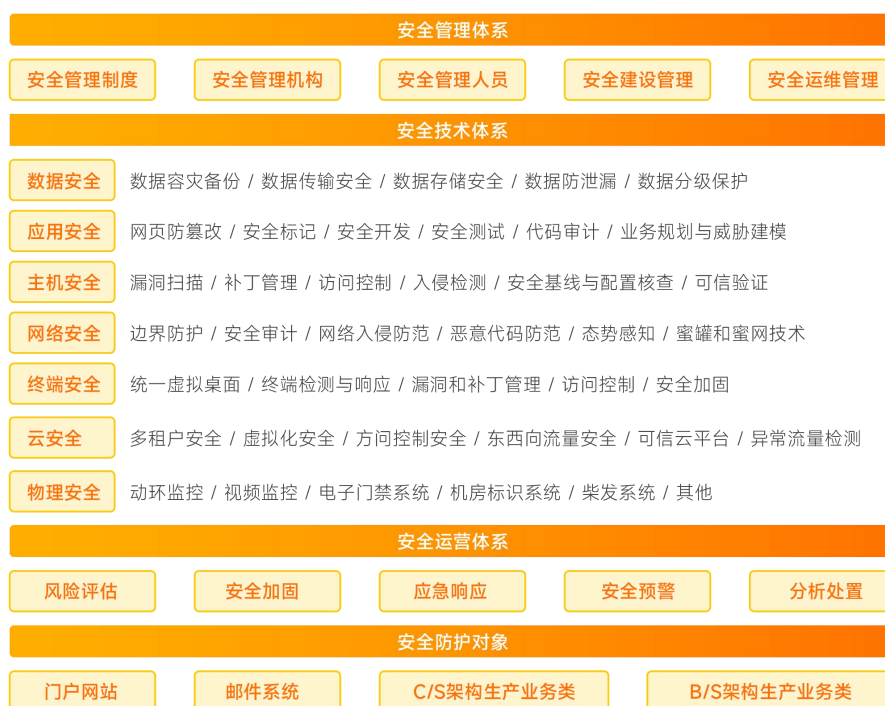
## 4.2 大数据安全体系的构建

尽管等级保护工作的推进有效保障了医院数据的信息安全，但要应对大数据时代的数据资产化及可能的数据流通，医院必须进行符合全新需求的安全体系建设，才能避免数据相关问题带来的损失。

(1) 数据存储和处理保护需求。按照国家及卫健委出台的各类政策及法规要求，医院均会采用“双活”及容灾备份方式进行数据的物理安全性保护和业务连续性保护，但是考虑到各类医疗数据自身的业务特点、数据特征、数据泄露途径等多方面因素，不可能采用一个产品或一种技术手段实现数据存储保护；医院相对开放的网络环境和脆弱的安全防御措施，极大地增加了诊疗数据可能的盗访、破坏，甚至遭遇勒索，对患者安全、医疗安全及质量、医院声誉、患者隐私等构成严重威胁，甚至酿成社会事故。

(2) 数据传输和交换保护需求。当前我国有关医疗数据保护的处理、分析等核心技术仍需完善，医疗机构的关键信息基础设施或系统数据传输、交换技术严重依赖发达国家，数据保护技术能力有限；众多的数据交换业务，急剧上升的互联互通需求，各种政策要求的数据上报，远程医疗工作需要，势必对单体机构的医疗数据私密性保护形成挑战；医学教育、培训、科研等以数据驱动的医疗发展带来的需求，对数据脱敏，海量数据的传输、检索，与应用安全防护形成矛盾。

图表 18 新一代医院数据中心安全体系架构



数据来源：《新一代医院数据中心建设指导》

通过对医院数据安全的现状调研分析，摸清安全技术的需求，针对医疗大数据的数据保护需要关注多个维度，特别是数据的物理安全、逻辑安全、隐私安全，形成安全防护体系，才能有效实现数据保护。业务数据的安全按不同维度可以分为三类：数据的物理安全性、数据的逻辑安全性以及数据的隐私安全性。

### 4.3 数据安全的保障逻辑

业务数据的安全按不同维度可以分为三类：数据的物理安全性、数据的逻辑安全性以及数据的隐私安全性。

#### 4.3.1 保障医疗数据的物理安全性

物理数据的安全性是数据安全的基础，保障数据物理安全性可以通过“双活”、数据快照技术、容灾备份等技术手段实现。

按照国家出台的网络信息安全等级保护条例、“智慧医院”建设标准，以及医院信息化互联互通成熟度测评标准等政策，需从数据的物理安全性以及业务的连续性角度，三级医院数据物理安全需要做到“4321”四项标准。其中，“4”代表相同的数据应该保存4份，其中2份数据为本地数据中心“双活”数据平台数据，1份为本地数据备份，另一份为异地数据实时保护数据；“3”代表为应该采用3种不同的技术手段实现数据的保护，例如选择常见的各种数据保

护技术手段：存储“双活”技术、数据快照技术、本地数据备份、异地备份技术、CDP 连续数据保护等；“2”代表至少保证医院有 2 个数据中心在同时运行，这样可以实现数据的应用“双活”；“1”代表为所有的数据的物理安全性应该统一管理，防止出现数据异步、割裂等。医院应尽量满足上述要求，以实现全面有效的数据物理保护。

### 4.3.2 保障医疗数据的逻辑安全性

逻辑存储安全是数据存储安全中的一个重要部分，医院需要对数据的逻辑存储（如认证鉴权、访问控制、日志管理、安全配置等）进行管理，保障数据存储安全。

有效的医疗数据逻辑存储安全手段需要针对不同特征的业务系统采用不同的保护手段来实现，按业务系统场景分为核心业务数据库、虚拟化系统应用以及海量医疗数据。

一、核心业务系统数据库。例如 HIS、PACS、EMR、LIS 等核心业务系统，数据库容量不大，均可采用数据库自身的备份接口，通过自身备份功能或者第三方商业备份软件完成数据库备份，按时间节点和数据库性能，实现数据完整备份和增量备份，结合简便性和易用性更高的 FC-SAN 存储，同时将备份任务和备份存储空间做好故障切换，提升备份速度。

医院亦也可采用其他高性能、具备源端、变长数据消重的专业备份设备，通过万兆网络以 NAS 的方式向数据库集群提供备份空间。

二、虚拟化方式运行的应用系统。针对以虚拟化方式运行的应用系统，各个虚拟化平台都提供相对应的备份接口，其实现方式与第三方商业备份软件思路类似，通过虚拟化平台快照来进行虚拟机的整机备份。

三、在线快照。传统的备份方案面对海量数据，其处理能力和资源耗用太大，技术上采用在线快照的方式来实现数据的逻辑安全性保护。目前实现快照有两种方法：重定向写快照技术 ROW 和拷贝写快照技术 COW。

四、COW 在进行快照操作之前，不会占用任何的存储资源，也不会影响系统性能。但该方式降低源数据卷的写性能，且如果主机写入数据频繁，那么这种方式将非常消耗 I/O。ROW 不会降低源数据卷的写性能，解决了 COW 重复写入导致的性能问题，在分布式存储上，ROW 的连续读写性能比 COW 好，但 ROW 的快照卷数据映射表保存的是源数据卷的原始副本，而源数据卷数据指针表保存的则是更新后的副本。因此，当创建了多个快照时，会产生一个快照链，使原始数据的访问快照卷和源数据卷数据的追踪以及快照的删除将变得异常复杂，在恢复快照时会不断地合并快照文件，造成较大的系统开销。

### 4.3.3 保障医疗数据的隐私安全性

在对医院内部的数据进行私密安全性建设时，必须构建完整的敏感数据保护体系，在医院内

部做好两个基础，管好三类人员与各类系统。

(1) 做好两个基础。指做好敏感数据分类分级机制和以三权分立为基础的制度保障。从总体角度看，数据库环境敏感度远远大于主机系统和网络层面，因此首先应该将敏感数据管理聚焦到数据库层级上，考虑以表为单位的敏感数据分类，把不同的表归类到不同的敏感数据集合中去；数据管理类似财务管理，必须按职权区分权、责、利三个角色，通过互相监督降低问题发生的概率，对所需要保护的敏感数据，通过设立三权分立机制保障敏感数据的安全性。

(2) 管理好三类人员。三类人员主要是指业务操作人员、运维和开发队伍、入侵者，前两者是系统应用、维护的主体，入侵者则带有典型的攻击目标特性。不同的人员，访问数据的行为特征也不尽相同。

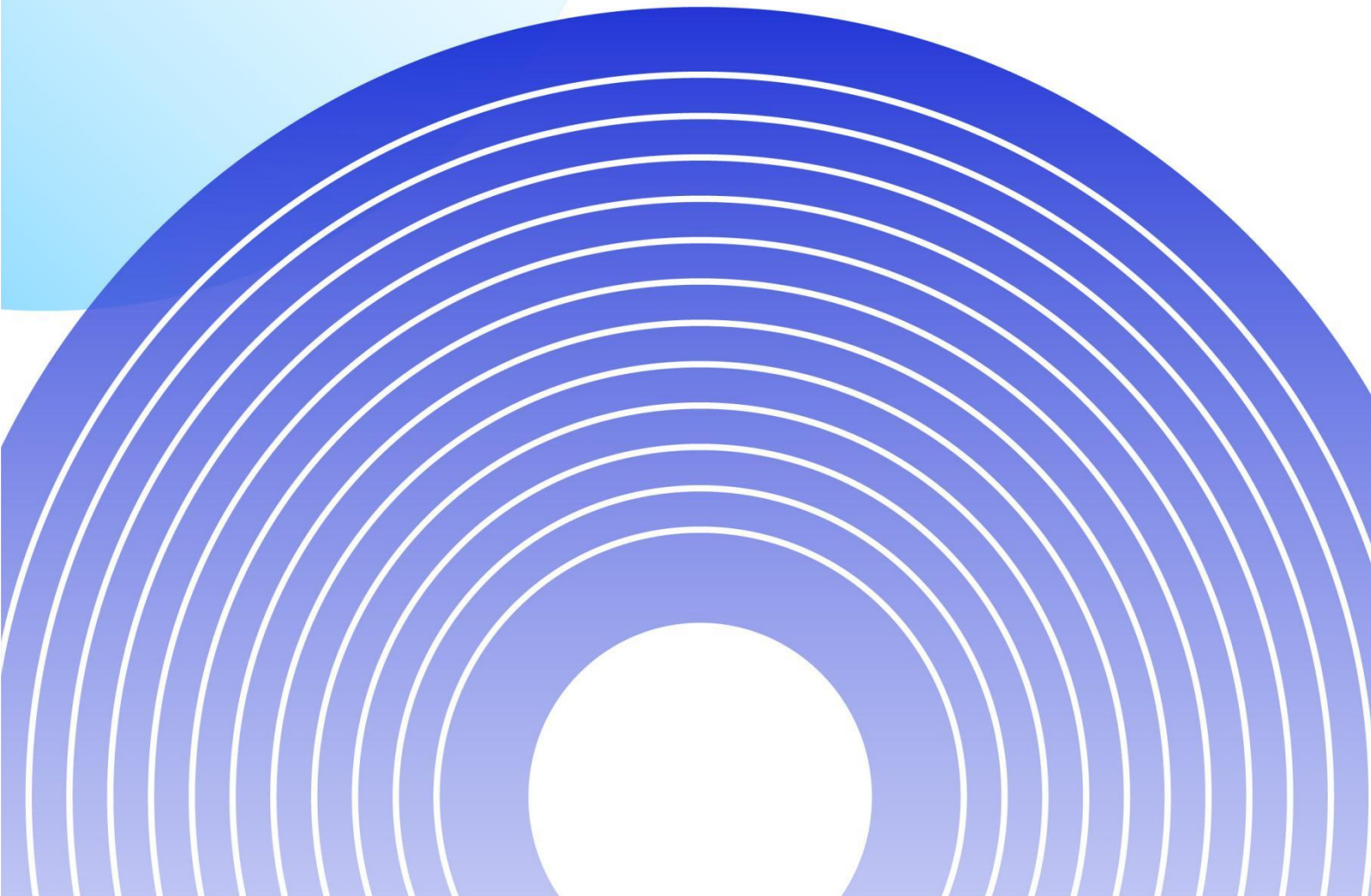
(3) 管理好常见 HIS、PACS、EMR、RIS 为代表的核心业务系统与其他一般医院业务系统。以基因测序、人工智能辅助影像诊断为代表的新兴医疗应用系统以及从事运维开发管理的运维开发类工具。要严格控制应用程序的访问和操作行为，通过从访问工具的合法性上管理好四类应用，同时根据医院自身情况及 IT 投入等多方面因素，将这四类应用数据进行不同级别的数据加密、脱敏。

# BIG DATA

## 第五章

前沿讨论

大语言模型的构建与数据交易的可能



## 5.1 讨论一：LLM 对于多模态大数据治理的潜在影响

尽管 NLP 的发展有力推进了智慧医院的建设，但落在具体场景中，如自动书写病案、智能问诊、智能随访等，该技术仍然没有脱离关键词映射数据库的逻辑，没有能够真正做到智慧智能。

大语言模型（LLM）的出现能够一定程度解决现有技术面临的智能程度不够问题。在分析文本类信息时，LLM 不仅能够从大量给定信息中找到任务需要的关键项，还能对未知信息进行预设，综合上下文做出推理。

相较于千亿级参数的通用大模型，医疗中文本类大模型的参数可控制于 100 万以内，包含文本与多模态影像的大模型参数可控制于 500 万内，因而非头部互联网公司也能参与医疗 LLM 的建设。

图表 19 医疗大模型企业图谱（截至 2023 年 9 月 20 日）



数据来源：蛋壳研究院

不过，从概念到落地，现阶段的医疗 LLM 仍需解决两个问题。

一是部署。当企业将大模型部署至医院时，需要医院购置相应 GPU 驱动模型运行。通常而言，服务一个科室的应用需要的 GPU 成本在数千元左右，但要负担全院需求，医院可能划分百万元级的成本购置芯片，因此，要推动 LLM 应用大规模落地，一方面需要推动医院主动部署 LLM 运行环境，另一方面需要企业方优化模型，尽可能降低医院在基础设施方面付

出的成本。

二是应用。目前基于 LLM 构建的智能应用仍然没有脱离传统医疗 IT 应用的范畴，如病案质控、智能问诊等，企业需要围绕医院需求构造“杀手级”应用，唤起医院购置的 LLM 的需求，进而实现 LLM 的规模落地。

构建 LLM 需要的成本不菲，且需大量医疗数据，因而竞争仍存在于头部医疗 IT 企业与互联网企业间。由于 LLM 应用需置于医疗信息系统中，因而非医疗 IT 公司只能使用外挂的方式进入 IT 环境，操作流畅性受限，相较之下，拥有医疗信息管理系统的医疗 IT 公司占有优势。同时，LLM 对于医院的架构要求严苛，能够支持 AI 应用的智能架构将比传统 EA 企业架构更好展现 LLM 的能力界限。

图表 20 传统企业架构与 AI 加持的架构的对比



数据来源：卫宁健康

## 5.2 讨论二：应用级医疗大数据的交易可能

就目前数据生产要素化形式看，临床数据可能是最具价值且能最快完成资产化的数据类型，但由于相关模式缺乏政策支持，仅能借鉴海外经验，探索应用级医疗大数据的交易可能。

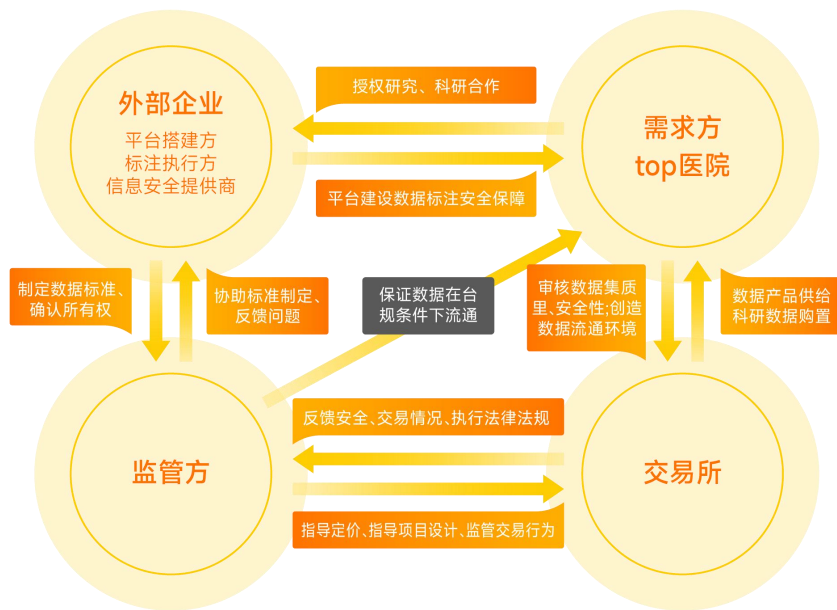
杭州国际数字交易联盟、浙江垦丁律师事务所研究显示：国外对于公共数据的利用大多采取类知识产权的“授权许可”模式在欧盟《开放数据和公共部门信息再利用指令》中，欧盟对公共数据采取使用许可证的授权方式，即公共部门通过行政许可机制授予企业对于公共数据的使用权。通过公共部门与企业达成的数据开放协议，作为企业取得公共数据使用权的法律依据，以及企业后续开发与使用公共数据的行为规范。在公共数据开发利用的收费问题上，欧盟采用多种方式的收费原则。一是引导公共数据低成本开放并对开放数据免费，但允许收

取复制、提供和传播公共数据产生的成本，以及消除个人隐私、为保护商业秘密而采取措施而产生的成本；二是对公共事业单位和需要自主创收的公共部门，按照客观、透明、可核查的标准进行数据定价，欧盟成员国在官网上公布此类公共部门名单，三是对高校图书馆、档案馆、博物馆等机构允许合理的投资回报以保障发展。

美国依据《著作权法》对具有版权或相关权利的公共数据库，采取知识共享 COO 许可，开放数据库许可等方式，授权用户用以商业或非商业目的的开发利用。英国依照《政府许可框架》《自由保护法案》等，对受版权或数据库权利保护的数据采用开放政府许可，允许用以商业或非商业的免费复制、发布、分发、传输及改编数据；针对超过《公共部门信息再利用条例》规定范围的数据再利用设置收费许可。

相较于发达国家，国内数据资产化程度缓慢，流通要素缺失。要构建完善的医疗数据交易体系，应协同监管机构、医院、交易所、第三方服务机构合作，共建医疗数据流通体系。

图表 21 理想状态下的数据要素平台闭环



数据来源：蛋壳研究院

### 1. 构建以患者个人信息为中心的全流程医疗数据安全防护体系

国标委 2019 年发布的《个人信息安全工程指南（征求意见稿）》中对医疗健康行业信息安全风险考虑要点，认为医疗健康行业是一个特殊的行业，其特殊性在于它以“人”为研究对象，所有医疗行为及其结果都以获取个人信息为基础，因此其信息安全风险非常高。

目前的患者个人信息安全防护，大多只是注重脱敏，进行匿名化、去标识化及加密等处理。由于医疗机构收集患者个人信息的特殊性，常规的“去标识化”无法完全保证个人信息不被复



原，还需要在专业医学人士、信息安全人员共同努力下对个人信息进行“去标识化”处理。在无法准确判断时，对于已进行“去标识化”处理的患者个人信息应采取严格的技术保护措施。

而对于专科医院，比如儿童医院，还需要遵从儿童个人信息保护的相关法律法规规定及国家标准，比如国家网信办 2019 年发布的《儿童个人信息网络保护规定》；《个保法（草案二审）》第十五条指出，处理十四周岁未成年个人信息的，应当征求未成年人父母或其他监护人同意；国家标准《信息安全技术 人脸识别数据安全要求（征求意见稿）》提出，在开展人脸验证或人脸辨识时，原则上不应使用人脸识别方式对不满十四周岁未成年人进行身份识别。

构建“以患者为中心”的个人信息风险评估和防护体系，覆盖个人信息收集、存储、使用、加工、传输、提供、公开等多个环节，有助于医疗机构促进和规范当前数据合规工作的推进和管理。

## 2. 建立完善的数据合规组织保障体系，设立数据保护官岗位

目前一些大型的医疗机构都设置了信息处及法务处，但其岗位设置和管理制度并不完善，组织机构职能割裂，难以更好地融合，随着医疗数据监管的日趋严格，将不能支撑医疗机构复杂的数据合规管理体系。

《数据安全法》规定：“重要数据的处理者还应当明确数据安全负责人和管理机构，落实数据安全保护责任”。而我国的个人信息保护法从第一稿开始即设立了“个人信息保护负责人”的岗位，并沿用至二审稿。

为充分保障医疗数据的安全，更好地进行数据合规体系建设，建议医院机构组织成立医疗大数据安全领导小组，在组织架构中至少还应该包括医疗大数据安全委员会和医疗大数据安全工作办公室，以确保做好医疗数据安全管理工作。其中医疗大数据安全委员会对医疗数据安全工作全面负责，讨论决定医疗数据安全重大事项；医疗大数据安全办公室指定专人，设立数据保护官（DPO）岗位，负责医疗数据安全日常工作。

数据保护官负责制定医疗数据安全和个人信息保护策略、风险评估方案、合规评估方案、风险处置方案和应急处置方案；负责建立数据安全和个人信息保护相关规章制度；负责人员的数据安全和个人信息保护方面的教育和培训；审计医疗数据使用和应用情况等。

## 3. 引入第三方服务，提高数据合规建设的能力

大多数医疗机构在体制上属于事业单位，人员主观能动性欠缺，医疗机构建立的内部安全管理制度和设计是否合法、合规、合理是个问题。引入第三方服务，比如律师、技术专家、咨询顾问等，是提高对医疗机构数据合规能力建设的有效方式。这些人员应在数据安全和个人信息保护领域有一定的工作经验，具备相关知识，熟悉法律、行政法规、规章及政策规则，

# 第六章

企业价值

助推医疗信息化进程的重要力量

BIG DATA

## 沛心科技

沛心科技成立于 2021 年 9 月，由苏州市产业技术研究院介入医疗技术研究所孵化，主要从事医学影像人工智能辅助决策类产品的研发和生产，现有产品包括 CT、超声、磁共振等医学影像自动标测及图像融合的 AI 系统、辅助诊断和手术规划 AI 平台系统。其中，沛心科技的“CardioVerse 心脏瓣膜分析系统”已进入商业化。

### 专注瓣膜手术，沛心科技以 AI 辅助风险评估

心脏 CT 检查是 TAVR 术前必不可少的环节。“经导管主动脉瓣置换术（TAVR）流程优化专家共识 2022 版”明确指出，术前应仔细阅读超声心电图、计算机断层扫描影像（CT）评估瓣环大小、主动脉弓角度、冠状动脉开口高低、主动脉瓣膜钙化情况。

目前国内瓣膜手术的 CT 术前规划，多由医生与瓣膜企业的技术人员共同完成，对技术人员较为依赖。国外虽有成熟的术前辅助产品，但购买价格昂贵，单机授权动辄数十万元。从这一需求出发，沛心科技自主研发的心脏瓣膜分析系统“智心”CardioVerse 便是为了解决上述问题。

具体而言，“智心”CardioVerse 是基于 CT 影像数据的瓣膜评估及手术规划系统，帮助术者进行手术风险评估，指导手术策略，保障 TAVR 手术安全性，提升手术成功率，助力 TAVR 术式下沉。与其他企业相比，沛心科技在研发及商业化的过程中具有其与众不同之处在于以下四点：

“智心”CardioVerse 在研发迭代的过程中，全程获取一线瓣膜公司和临床医生真实反馈，在反馈基础上高频率迭代，使产品更符合中国医生的操作习惯；

该系统对电脑操作系统和配置的兼容性强，可推动介入瓣膜手术在基层医院的普及，造福更多瓣膜病患者；

交互界面友好，兼顾不同医生的诉求，拥有简单模式和进阶模式，既能帮助基层医生快速上手应用，也能满足资深医生灵活的科研需求；

可结合嵌入式硬件系统（CT、DSA、手术机器人等）匹配术中导航系统，利用多模态影像融合技术优化瓣膜手术流程。

目前，沛心科技的 TAVR 模块已经进入商业化进程，全国已装机授权数百台，装机医院近百家。此外，沛心科技还与多家瓣膜公司合作，初步实现该领域的国产替代。

### 完全自主知识产权支撑，AI 能力向神经介入扩充

据沛心科技表示，该公司拥有完全自主知识产权的底层代码库，产品的延展性、复用性及移

植性极高。同时公司拥有一支经验丰富和年轻化相结合的研发团队，产品迭代速度极其快。“产品好用、团队反应速度极快、迭代速度惊人”是公司获得最多的评价。

在主动脉瓣模块成功应用的基础上，沛心科技已布局二尖瓣、三尖瓣、肺动脉瓣及结构性心脏的其它领域。TAVR 模块商业化后，沛心科技加快了研发和产品定型速度，并在算法优化的基础上应用图像融合技术拓展功能至瓣膜病辅助诊断及早期筛查。

未来，公司将扩展神经介入领域及影像云系统的应用，布局“智影”和“智脑”产品矩阵，借助人工智能进一步解决神经介入诊疗流程中的痛点。

## 卫和医学

北京卫和医学科技有限公司是一家以健康危险因素大数据为基础的科技公司。基于权威的医疗健康行业知识引擎，为健康管理公司、数字化医疗行业等提供针个性化健康管理服务及医学知识图谱。

### 动态数据库创新，卫和医学紧随智能化时代知识库需求

伴随智能化时代的推进，市场对于医学知识库的需求不断转变，已由单一静态知识库需求转为动态需求。这意味着，知识库提供方不仅需要满足使用方对于知识节点的单体描述进行查询与搜索，还需要根据单一知识点能够获取该知识领域相关信息，并且能够定位到该知识点的多维描述，让用户一目了然看懂整个知识结构。

为了应对上述需求，卫和医学打造的动态医学知识库可以对每一个数据节点及知识点都有很好的描述与关联。并且根据疾病症状、体征、鉴别诊断、药品、检查、检验等能很清晰快速地定位。此外，该动态知识库也是一类基于人工智能技术的医学知识服务系统，它可以根据患者的病史、症状、检查结果等综合诊疗信息，自动生成相应的诊疗、护理及康复等符合临床需要的方案。

具体而言，动态医疗知识库的特点包括实时更新、自动化处理、可扩展性三点。实时更新意味着动态医疗知识库可以随时更新最新的医学知识和研究成果，以保证其准确性和实用性。自动化处理则表现为动态医疗知识库可以通过机器自动化实现快速补充及完善，减少人为错误。可扩展性体现为动态医疗知识库可以根据需要进行扩展，以适应不断变化的医学需求。

### 可视化解释动态医学内核，卫和医学为医学 AI 构建强硬基底

由于卫和医学知识库拥有可视化解释动态医学内核，可辅助医疗 AI 公司进行辅助诊断、临床路径、医疗知识点结构、诊疗方法与环境、手术检验检查、医药品知识库等研究，打造 CDSS、AI 辅助诊断、动态风险识别、AI 机器人、康养健康监护等应用。

LLM 方面，卫和医学已将其知识库产品应用在家居产品方面，如智能马桶，使得数十条检测数据不再是看不懂的数字 1、2、3 或是阴性、阳性等只有医生才能看懂的结果，而是转变成浅显、通俗易懂的百姓语言。可穿戴设备方面，卫和医学推出了卫和护心衣，可实时识别用户生命体征综合数据，如心脏年龄、压力指数、各种心律失常及心脏病、心源性猝死等心脏健康相关医学知识内核，助力百姓远程医疗与养老健康等行业广泛应用。

到目前为止，卫和医学已经为多家行业知名医疗 AI 企业提供动态知识库支撑，助力国内近十款 AI 大模型的生成与发布。未来，卫和医学将会继续耕作医学知识海洋中的探索与研究，更加紧密的与专家医生联合加强审核医学内容严谨性、科学性、安全性，积极推进知识节点

的更加精细化与知识更新。

衍生品方面，卫和医学将进一步围绕百姓生活打造健康解决方案，将医疗知识的精准推送到合适的场景中，为百姓的健康预防与疾病干预尽一份力。

## 上药云集团

### 探索数字化医药与创新支付新空间，广州益药·云药房打出标杆案例

历经三年快速升级发展，位于广州医药物流园区内的上药云健康广州“益药·云药房”经营面积超过 1000 平方米，药品陈列空间达 200 平方米；主要合作伙伴直线距离均在 1km 以内，以创造高效运输模式保障药品按时供给。同时，药房设有 7\*12 小时客户服务呼叫中心，可承担各大医疗机构延伸处方以及互联网医院电子处方药品供应和配送需求，并可满足保险企业、大客户健康以及患者线上线下对于药品的用药需求。

截至目前，广州“益药·云药房”已成为专业特药直付服务供应商，为保险企业、大客户健康管理药品提供服务 45 万+人次；通过 B2B2C 药品销售新模式完成日均药品服务 4000 单。通过创新支付、商保直付、惠民保准入的特药直付模式，以及门诊险、慢病组合包、企业防疫包的普慢药增量模式，不断突破患者的支付困境，加速药品可及、推进渠道多元。

与此同时，通过全面实现药品供应“广覆盖+可及性”的专业化药事服务，广州“益药·云药房”充分发挥集约优势，为业务全链条拓展、持续探索数字化与创新支付新空间打下坚实基础。

### 大数据支撑，上药云健康助力药企精准营销和临床研究

上药云健康的药企数字化服务业务是基于公司广泛的专业药房服务网络和丰富的创新药全生命周期服务经验开发的创新合作模式。在创新药的研发阶段，上药云健康是中国最具规模体量的临床实验用药供应商，赋能 80+家国内外厂家，管理拥有 265+个在研项目，临床试验基地覆盖 90+座城市。同时，上药云健康的患者招募业务依托数字化创新基因，通过线上线下多元化患者招募方式，加速创新药的研发进程。

在创新药的上市阶段，通过面向药企、患者和药师的解决方案为药企客户赋能，如：DDI 数据直连、进销存数据管理、DTP 药房调研及洞察分析、患者首访一随访服务、患者 DoT 管理、药师招募注册、药师培训服务等。

作为全国首批连锁专业药房，上药云健康深耕创新药细分赛道多年，拥有广泛布局的专业药房体系、全生命周期配套服务、稳定成熟的创新模式设计，在帮助创新药企实现患者可及的基础上，也积累了宝贵的大数据和数据分析及应用机制。

目前，上药云健康就其经营的 6 大科组、130+个重大疾病和慢病病种，与协会组织、药学专家共同编撰完成相关癌种比如肺癌、乳腺癌等治疗药品药学服务规范，同时也是最早开设新特罕专科病种服务网络的专业药房，把专业药事服务贯穿到用药前、用药中、用药后的患者全周期服务旅程，挖掘患者群面临的多层次需求，不断探索增值服务模式，积累以患者为中心的全面、细致、完整的数据体系。

未来云健康将持续发挥“云”基因优势，在业务层面全面延伸对药企、对患者、对医疗机构和政府部门的全渠道合作和服务链条，在运营层面建设贯穿各个业务板块的前台—中台—后台运营和数据一体化生态，确保公司作为领先互联网+医疗科技平台，持续有效运转，建立行业标杆模式。



## 参考目录:

1. 数据资产价值实现研究报告 杭州国际数字交易联盟、浙江垦丁律师事务所
2. DIP 改革赋能三级医疗机构发展内涵的实践与应用研究 宋静 徐克平 米灿 刘远泉
3. 《肿瘤大数据与真实世界研究中国专家共识（2022 版）》 张艳桥 徐建明
4. 《全民健康信息化调查报告——区域卫生信息化与医院信息化(2021)》 国家卫生健康委统计信息中心
5. 《2021-2022 年度中国医院信息化状况调查报告》 CHIMA
6. 《新一代医院数据中心建设指导》 张雪高 胡建平
7. 薛万国：医院大数据中心建设的六个常见问题 HIT 专家网
8. 临床科研数据库系统的现状与未来 薛万国

免责声明：本报告的信息来源于已公开的资料和访谈，蛋壳研究院对信息的准确性、完整性或可靠性不作保证。本报告所载的资料、意见及推测仅反映蛋壳研究院于发布本报告当日的判断，过往表现不应作为日后的表现依据。在不同时期，蛋壳研究院可能发布与本报告所载资料、意见及推测不一致的报告。蛋壳研究院不保证本报告所含信息保持在最新状态。同时，蛋壳研究院对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

版权申明：本文档版权属于蛋壳研究院/北京蛋黄科技有限公司，未经许可擅用，蛋黄科技保留追究法律责任的权利。

蛋壳研究院（VBR）：蛋壳研究院关注全球医疗健康产业与信息技术相关的新兴趋势与创新科技。蛋壳研究院是医健产业创投界的战略伙伴，为创业者、投资人及战略规划者提供有前瞻性的趋势判断，洞察隐藏的商业逻辑，集合产业专家、资深观察者，尽可能给出我们客观理性的分析与建议。

研究人员：



赵泓维 高级研究员