



陸軍軍醫大學
第三附屬醫院

临床数据治理探索

黄 昊



重庆大坪医院信息科副主任
中国研究型医院学会医疗信息化分会常务理事
中国研究型医院学会医疗信息化分会
 智慧医疗专委会主任委员
中国医院协会信息管理专委会委员
中国卫生信息学会医院统计专委会委员
全军医学科委会卫生信息学专委会委员
重庆市医学会信息管理专委会副主任委员
重庆市医院协会信息管理专委会副主任委员





欢迎常交流，共同学习
构建医疗信息化知识体系



微博



微信公众号



从IT到DT, 医疗信息化正当时





时代特点

1

信息化需求强烈

2

数据价值被认同

3

医学研究依赖加强

4

IOT在临床开始规模化使用



时代特点

5

整合成为常规

6

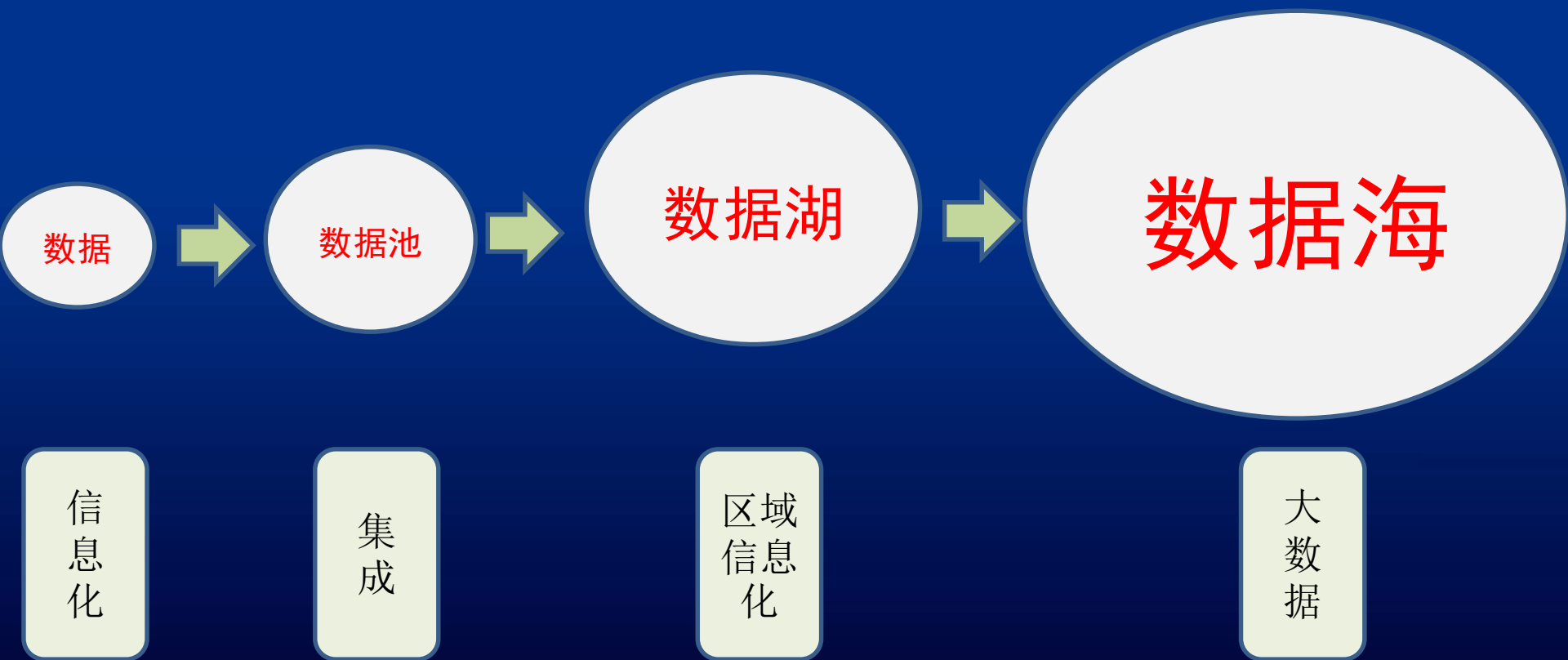
管理更加严格

7

信息安全纳入医疗质量管理



- 十九大报告
- 健康中国2030
- 十三五卫生规划
- 单病种付费
- DRGs



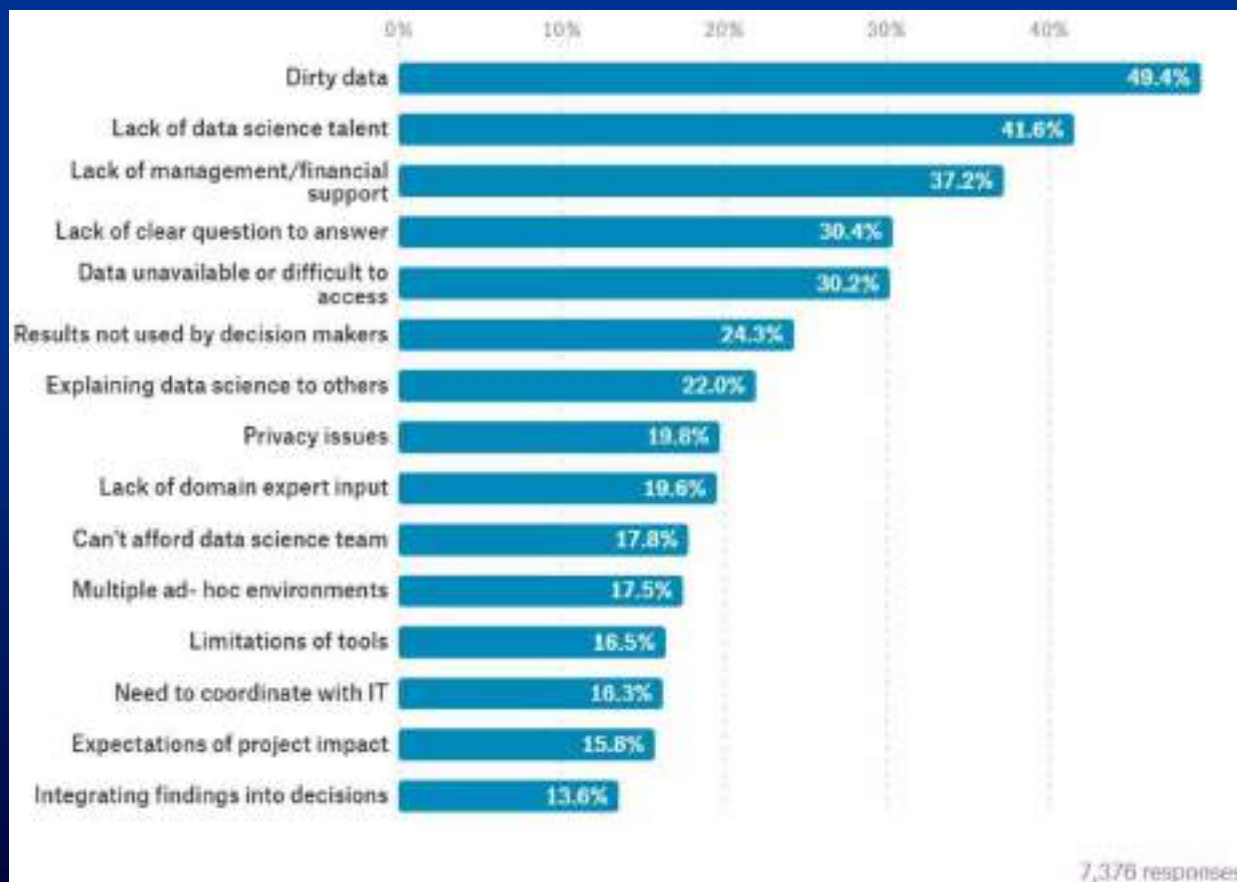


現 狀

- 1 数据质量堪忧
- 2 关心业务甚于数据
- 3 数据标准用的更少
- 4 数据安全更被忽视



数据质量堪忧

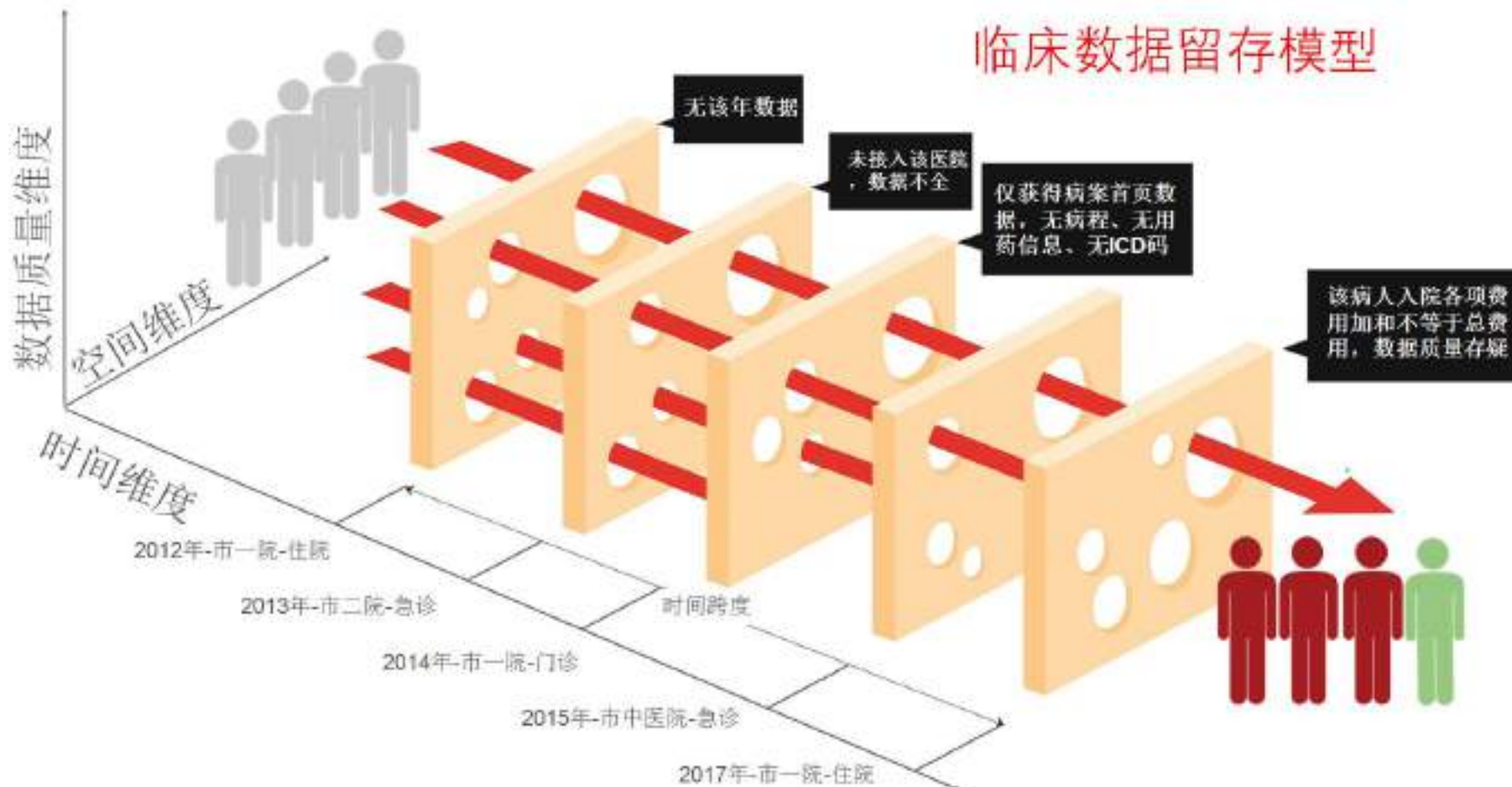


大数据时代的现状



临床数据面临的问题

- 安全威胁
- 数据质量不高
- 数据标准性差
- 数据准确性差
- 数据滥用





数据标准化不足

数据标准化，结构化程度低



ICD诊断
编码纠正



身份信息
纠正



药物名称
剂量统一



检查名称
单位统一



其他非自治信
息纠正

标准类别	适用范围	标准名	支持版本	级别
国家标准/ 国际标准	疾病名称	ICD-10	全国版2011,全国1.3, 北京 v6.01,上海2013,国标GB-T 14396-2016	必须使用
国家标准	电子病历	《病历质量考核评价标准实施细则考核表（手术科室终末病历部分）》 《病历质量考核评价标准实施细则考核表（病历首页部分）》, 《病历质量考核评价标准实施细则考核表（病历首页部分）》, 《病历质量考核评价标准实施细则考核表（病历首页部分）》	各地各版本	必须使用
国际标准	症状	SNOMED-CT, UMLS	UMLS-2017AA	建议使用
国际标准	实验室检测	LOINC	Version 2.61	建议使用
国际标准	药物	RxNorm	RxNorm_full_08072017	建议使用
国家标准	药物	国家食药监药品本位码	NA	建议使用



数据颗粒度层级

地区级
质控变量

区域化数据完整度、代表性

医院级
质控变量

单医院数据完整度、质量

单病人级质控变量

单病人数据完整度、自治程度

临床间接质控变量

临床数据类别标准化程度

临床分类变量级质控变量

临床指标数据完整度

临床单（全）变量级质控变量

单变量完整度、合理性

数据是否覆盖各级医院、专科医院、中医医院

包含门诊、住院各科室历年数据

数据包含单病人完整病历、诊断、检查、用药信息

再入院间隔时长应合理

肝功能项目数据全

合理取值范围

示例



临床数据生命周期

临床数据全生命周期





数据产生

- 软件开发人员错误、失误
- 录入人员错误、失误
- 采集传感器、芯片带来的采集错误
- 设备中普遍存在的后门现象
- 非可信用户产生



数据传输

- 网络质量
- 误接、错接引起
- 移动设备的失控
- 数据泄密
- 数据渗透

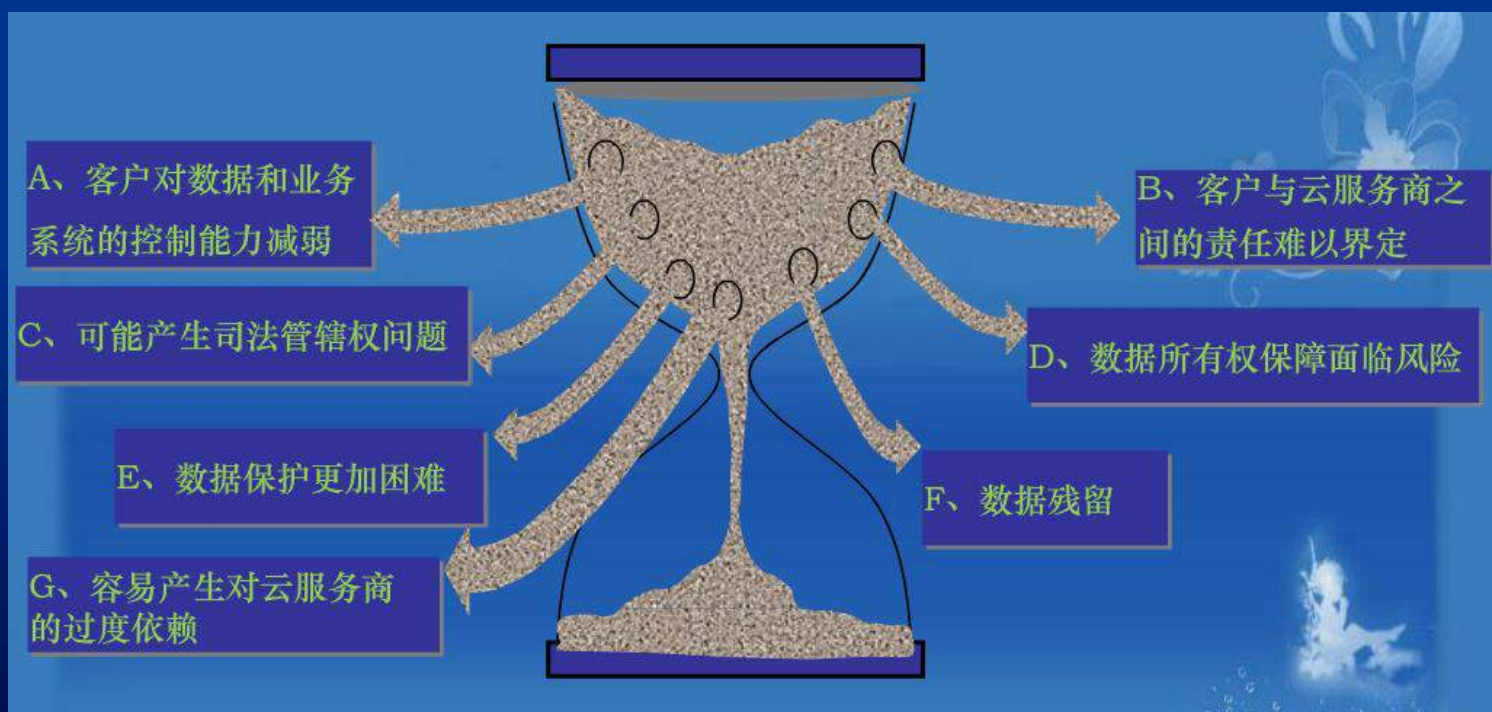


存 储

- 设备安全，配置是否合理
- 备份方式，备份策略
- 是否加密
- 密码是否强壮
- 数据共享的方式是否安全
- 云存储如何把控
- 被忽视的UPS



云服务的风险





分 析

- 分析模型是否可信，有没有经过大规模测试。
- 错误的呈现方式会让人产生误解
- 医学的认知依然有限，计算机的能力也还弱小
- 错误的数据会带来错误的结果的不断呈现



发 布

- 数据访问的权限控制
- 数据使用的责任
- 数据管理的防范措施
(合适的人合适时间合适内容)
- 患者隐私如何保护
- 医学伦理学合规的考虑



数据安全风险

美国医院信息勒索事件

2016年2月发生在美国洛杉矶的好莱坞长老会医院 (Hollywood Presbyterian Medical Center) 的网络安全事件给人们再次敲响了警钟。

2月5日
黑客控制
医院系统

10天内
医务人员
手工工作

2月15日
医院瘫痪
支付赎金



黑客：你们的电子病历系统已被我掌控
要想继续使用就给我付钱！

医院：...



电子病历：已被锁定，请输入密码...

医务人员：没有电子病历怎么开展工作啊！



医院：好吧，好吧，我支付赎金。



x40
=\$17000

赎金比特币：我是虚拟货币，来无影去无踪，休想查到我！



人的管理

顺丰数据失窃案

1 公司网站账号密码泄露

2016年8月26日，湖南省长沙市顺丰速递有限公司宋任宇，将个人所掌握的公司网站账号及密码出售他人，造成泄露的顺丰用户个人信息大量泄露。

2 编写恶毒程序批量下载客户信息

2016年7月东莞市石碣分部分管员陈洋伙同杨维保，利用杨维保编写的APP软件批量下载客户个人信息，一年间盗取10万余条。

3 内部人员大批量查询客户信息

2015年深圳多名顺丰客户接到诈骗电话，接举报，顺丰公司反查单号发现有员工一月间查询客户记录高达2.8万条，并多次利用他人账号查询信息。

4 通过购买内部办公系统地址、账号及密码，侵入系统盗取信息

2013年，杨某等人在互联网上购买顺丰内部网络系统登录地址、用户名及密码。通过无线网络侵入顺丰内部办公系统，批量下载客户个人信息及快递记录，非法贩卖从中牟利。截至案发，杨某等人共非法获取个人信息180万余条。

5 总部研发人员从数据库直接导出客户信息



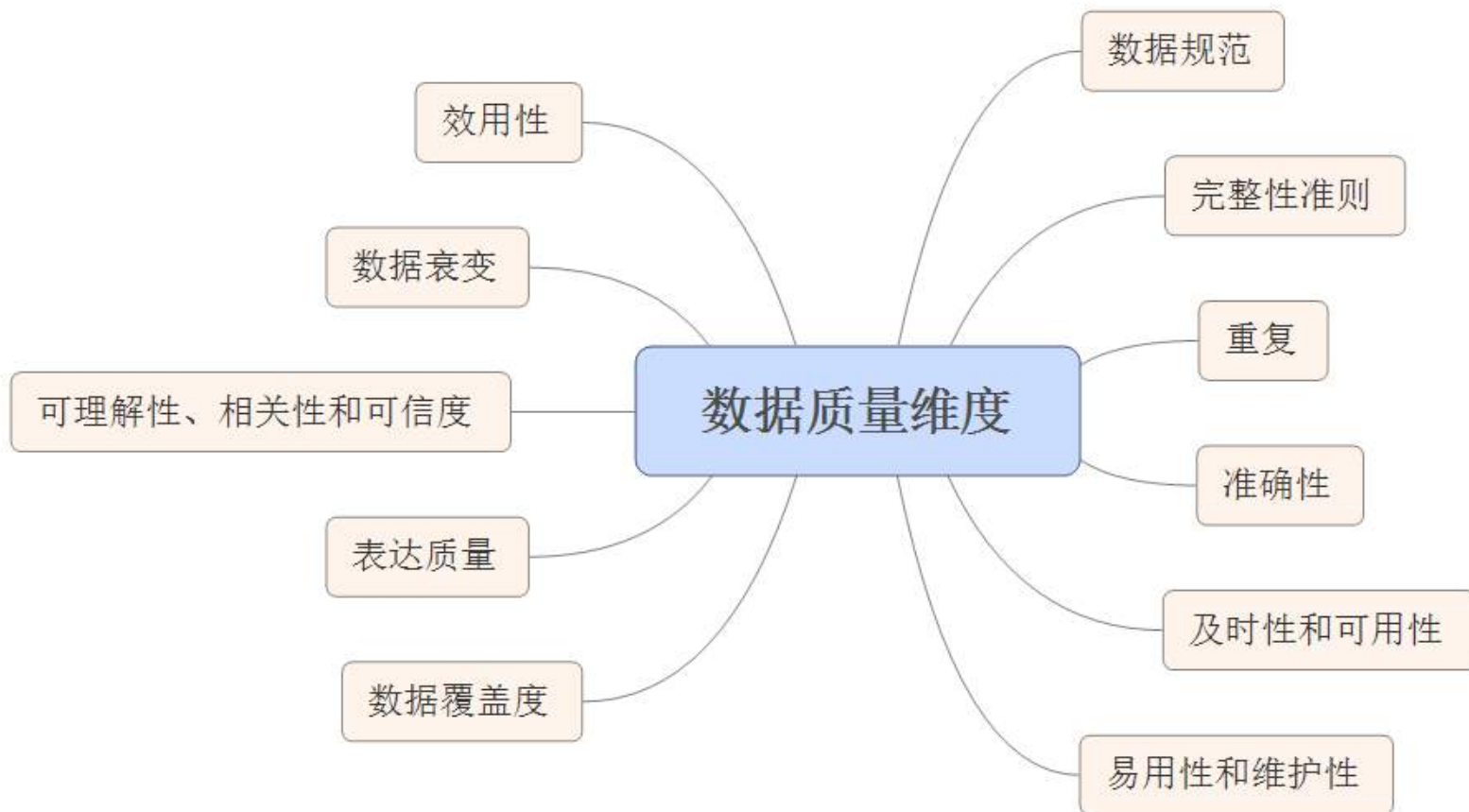
陸軍軍醫大學
第三附屬醫院



临床数据治理框架

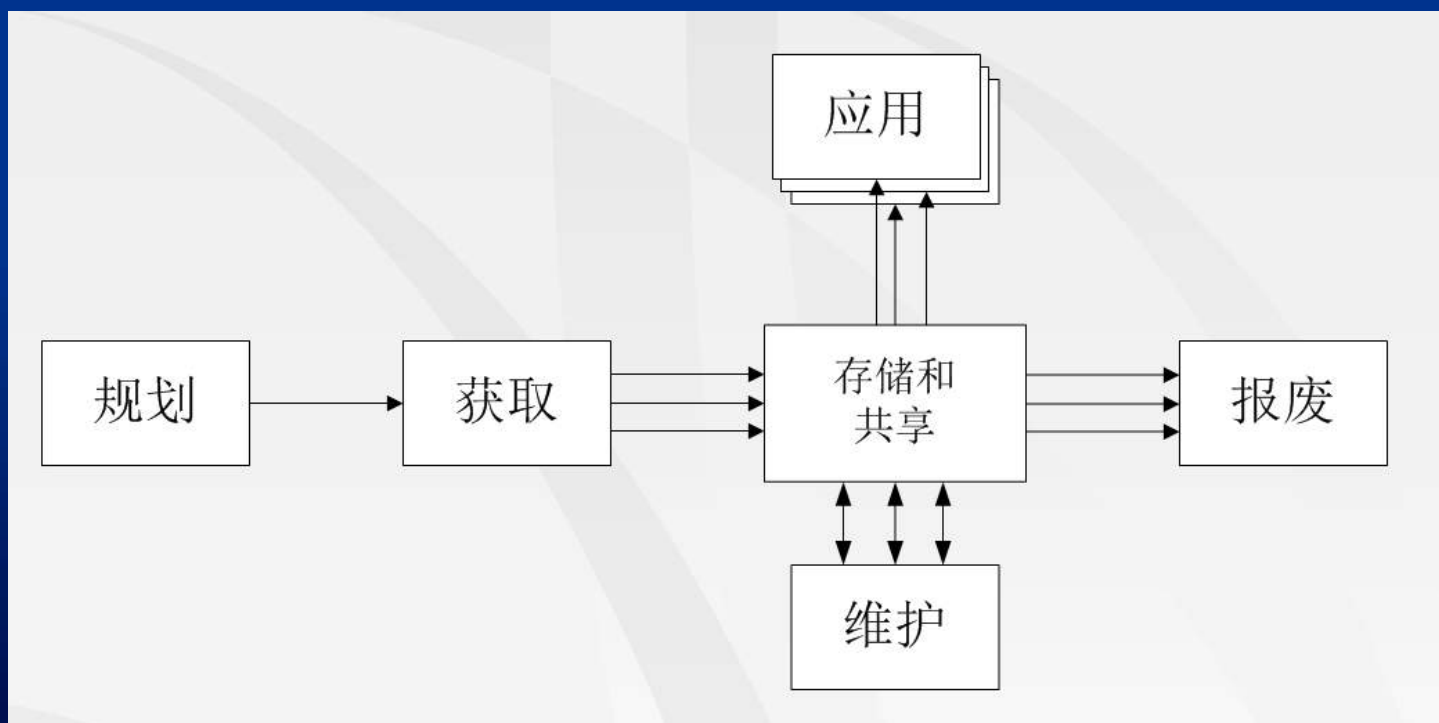


概念



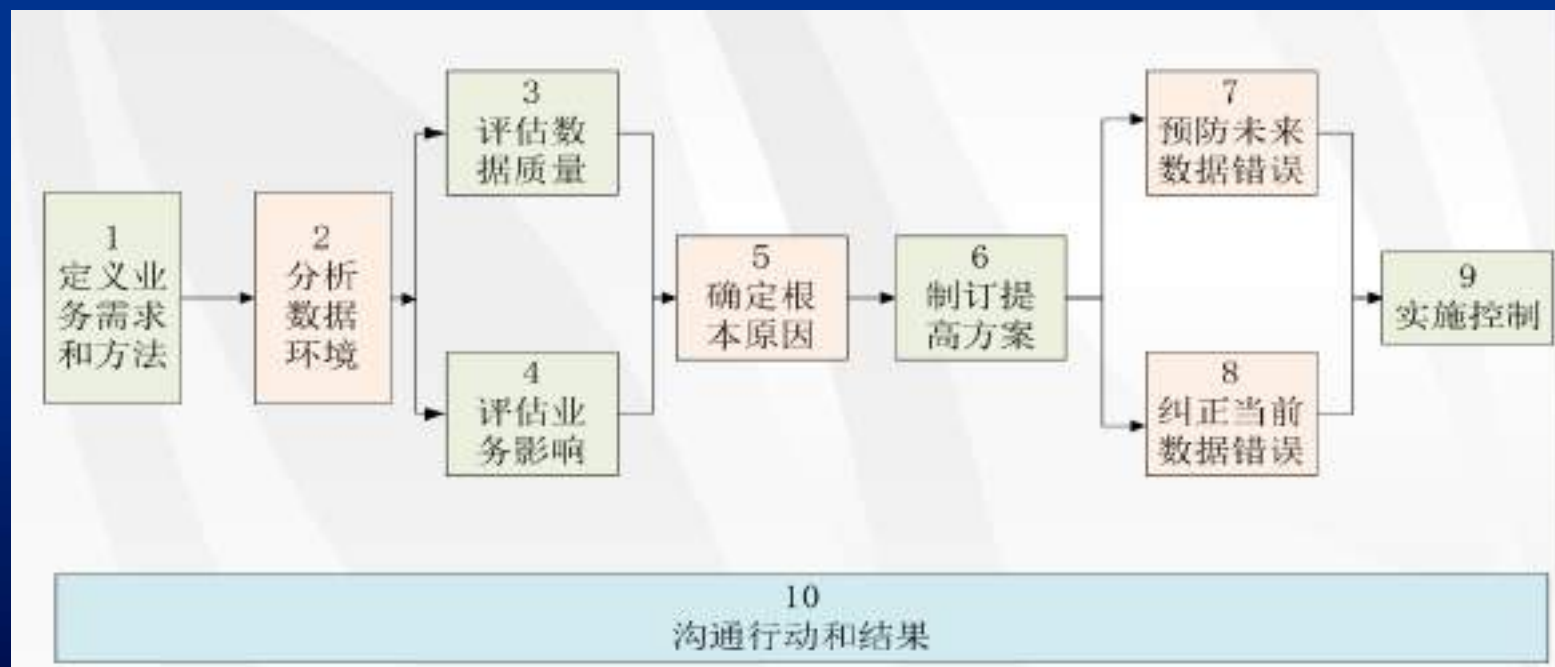


数据治理必须树立的理念





数据治理的十步法





信息和数据质量提高周期

评估

认知

行动

六西格玛的DMAIC

定义

测量

分析

改进

控制

1
定义业务需求
和方法

2
分析信息环境

3
评估数据质量

4
评估业务影响

5
确定根本原因

6
制订提高方案

7
预防未知数据错误

8
纠正当前数据错误

9
实施控制

10
沟通行动和结果



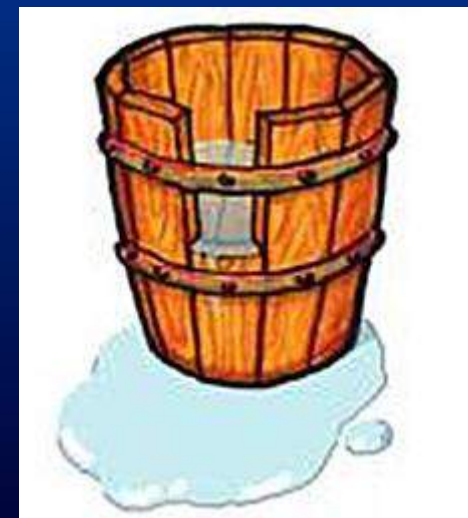
- 常见数据质量管理流程：
- 数据理解--了解和领会数据的含义和行为。
- 数据修复或改进--纠正或更新数据，如姓名和地址的更新，以及专门的查找替换功能。
- 数据监测--应用测试在数据流动或处于静止状态时检查数据，这可引起人们对新的异常现象或认为已解决的重点持续问题的注意。



- 数据剖析--列或内容剖析或分析--分析记录的每一列，检查数据集中的所有记录，列剖析将提供完整性/填充率，数据类型、大小/长度、唯一值列表和频率分布、式样、最大最小值范围。
- 内部表/内部文档/依赖关系剖析/分析--发现表和文档内属性之间的关系，能够发现表和文档的数据结构，函数依赖组件和数据结构的质量问题。
- 交叉表/交叉文档/冗余剖析/分析--比较表和文档之间的数据，确定重复或相同的数据值，确定重复值或标志外键。



- 数据管理幻化到日常工作
- 每年一次的数据质量评估
- 必要的经费投入
- 数据质量纳入绩效
- 文化氛围的营造





- （一）设置专门岗位
- （二）定期对从业人员进行教育、培训和技能考核；
- （三）对重要数据进行核查机制；
- （四）数据使用的合规性审查
- （五）法律、行政法规规定的其他义务。



防 范

- 识别风险
- 确定本院的风险管控成本
- 制订安全改进措施计划
- 设岗设人资金保证
- 寻找安全合作伙伴
- 全员教育培训



未 来

- 大数据洞悉数据质量问题
- 人工智能提升数据管理能力





- 信息资产可视化，分类识别、行为建模，特征分析，关联进行风险评估，安全态势感知。





人工智能提升安全能力

- 早发现，早处理
- 侦测和清除僵尸联网设备
- 阻止恶意软件运行
- 提高安全运营中心效率
- 风险评估更准确
- 构建安全态势模型



谢谢聆听



微博



微信公众号

数据利用从治理起步