



云计算 大数据

智能健康

健康

中科院上海健康科学研究所
上海交通大学医学院附属瑞金医院

姚志洪



中国梦



中国梦的核心是国富民强、人民幸福
保障人民的健康则是重中之重



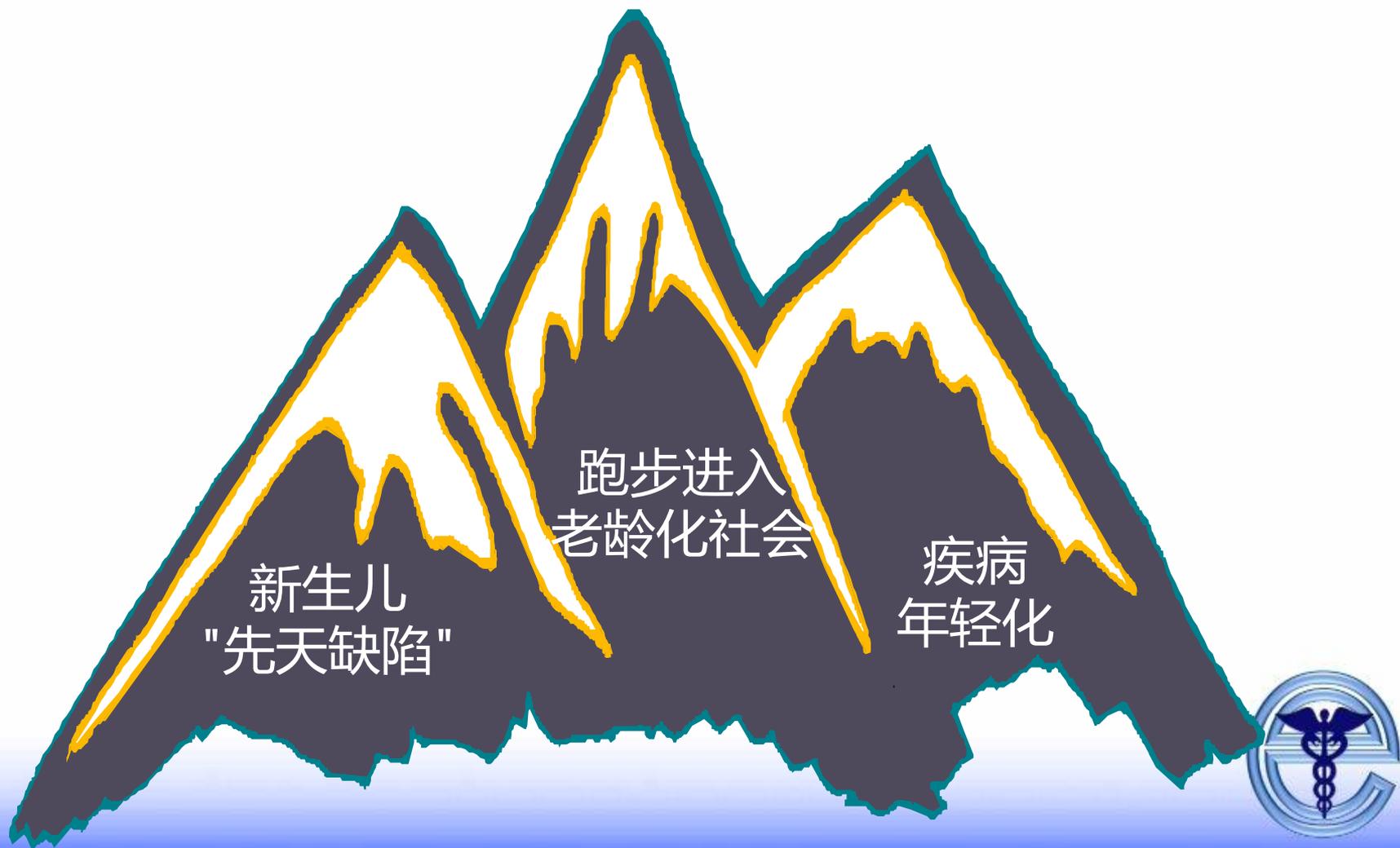
健康梦是倡导文明，关爱生命，促进健康，推动
“疾病医学”向“健康医学”转变，使人民不得病、
少得病、健康长寿



实现“中国梦”，要有健康的体魄
让健康托起“中国梦”



健康医疗预测报告



老龄化



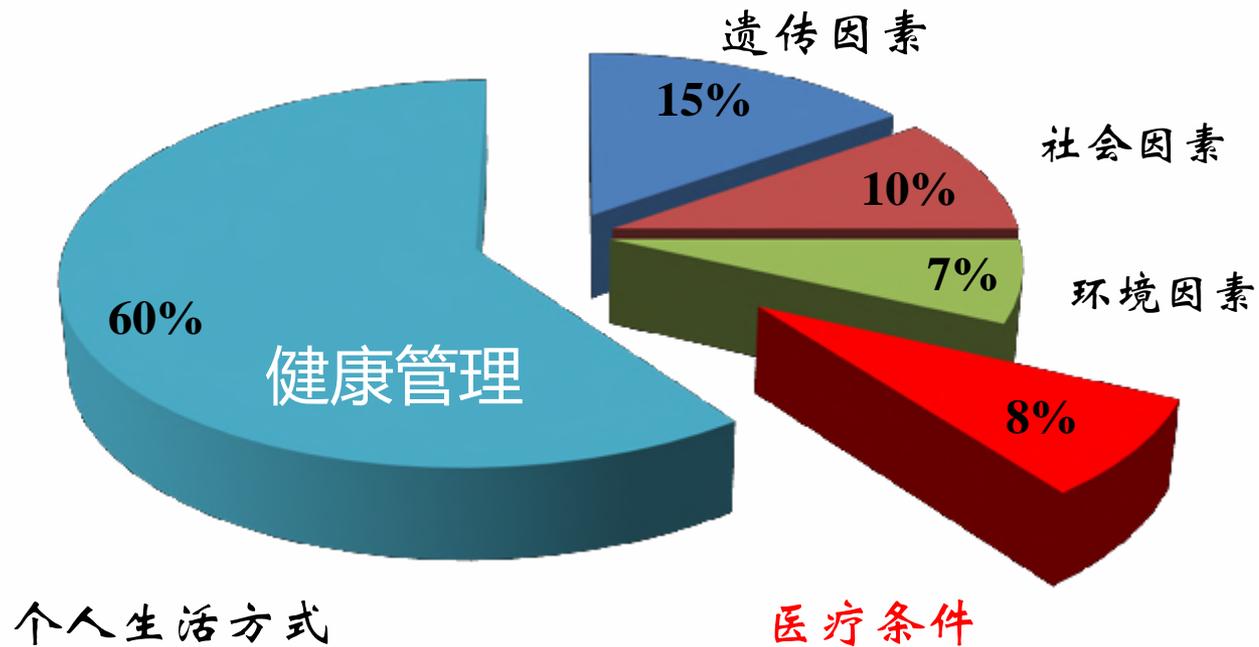
2013

需要耗费较多的医疗资源

2020



健康的影响因素



《世界卫生组织》



健康与经济



健康是人们期望的首位，是财富、基本人权、能力、智力和体力的基础



期望寿命提高10%，经济年增长率提高0.3% ~ 0.4%



1美元的健康投资可取得6美元的经济回报



信息化



我国医疗卫生资源严重短缺，远远满足不了快速增长的健康需求



健康产业关系到生老病死，是永盛不衰的大产业



云计算、大数据和智能健康，可以缓解供需矛盾，逐步做到健康服务公开、公平、公正

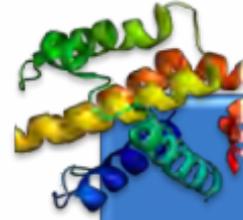


大数据来临



技术

- 云计算、大数据、物联网、移动互联网快速发展，促使健康数据正以前所未有的速度增长，健康大数据时代已经到来



资源

- 健康数据是一种基础性资源，管理和利用健康大数据为人类健康服务已经成为普遍关注的热点



价值



《大数据时代》作者维克托 “大数据比作新黄金”



邬贺铨院士 “大数据价值堪比石油，首要应用是智慧医疗”，大数据是新财富，是新型生产资料



健康信息化市场已成为一片蓝海，面对庞大而复杂的医疗和健康数据，大数据将会带来令人兴奋的回报。



预测

利用大数据提高效率和质量

对健康大数据进行分析

可创造超过3000
亿美元额外价值

3000亿

减少美国医疗保
健系统成本支出

200亿
到
300亿

减少美国医
疗保健支出

8%

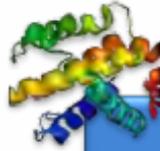


健康数据



临床医疗数据

- CT图像约150MB , 病理图像接近 5 GB
- 重症监护：24小时连续检测
- 结构化数据: 血压、心率、血氧
- 非结构化数据：心电图和呼吸等



高通量组学数据

- 个性化健康是以基因组测序为代表
- 组学数据
- 蛋白质组学
- 代谢组学
- 精准医疗



健康信息

- 个人健康信息采集是未来重要的来源
- 家庭健康检测
- 每天产生：体重、心率、血压、血糖、心电图、呼吸等



特点



- 规模性 (Volume)
- 多样性 (Variety)
- 易变性 (Variability)
- 价值性 (Value)
- 高速性 (Velocity)



多样性



○ 预定义数据类型、格式和结构的数据

ICD编码	ICD名称	拼音码	五笔码
A34.352	破伤风伴有流产或异位妊娠	PSFY	DPVVE
008	流产、异位妊娠和葡萄胎妊娠后的并发症	LCTVR	IURVV
008.052	异位妊娠手术后休克	YKSS	MYVVR
008.301	异位妊娠后休克	YKSH	MYVVR
008.802	异位妊娠术后盆腔残留	YKSS	MYVVS
010.051	原有特发性高血压并发于妊娠、分娩和产褥期	YTFX	DECVN
010.151	原有高血压性心脏病并发于妊娠、分娩和产褥期	YTGK	DEITD
010.201	妊娠合并肾性高血压	ESHBS	VVVUJ
010.351	原有高血压性心脏病的肾脏病并发于妊娠、分娩和产褥期	YTGK	DEITD
010.451	原有继发性高血压并发于妊娠、分娩和产褥期	YJFX	DECVN
010.901	妊娠合并高血压	ESHBG	VVVUJ

ICD编码、挂号、收费、药品

○ 可识别和解析的文本数据文件

```
--HL7-CDM-boundary
Content-Type: text/xml; charset="US-ASCII"
Content-ID: <10.12.4547.43>

... Base 64 of of Base CDM document, which contains
...
<ObservationMedia>
  <id root="10.12.4547.345"/>
  <value mediaType="image/jpeg">
    <reference value="camed_left_hand_image.jpeg">
      </reference>
    </value>
  </ObservationMedia>
```

XML格式
电子病历

○ 没有固定结构的数据文件



检查报告

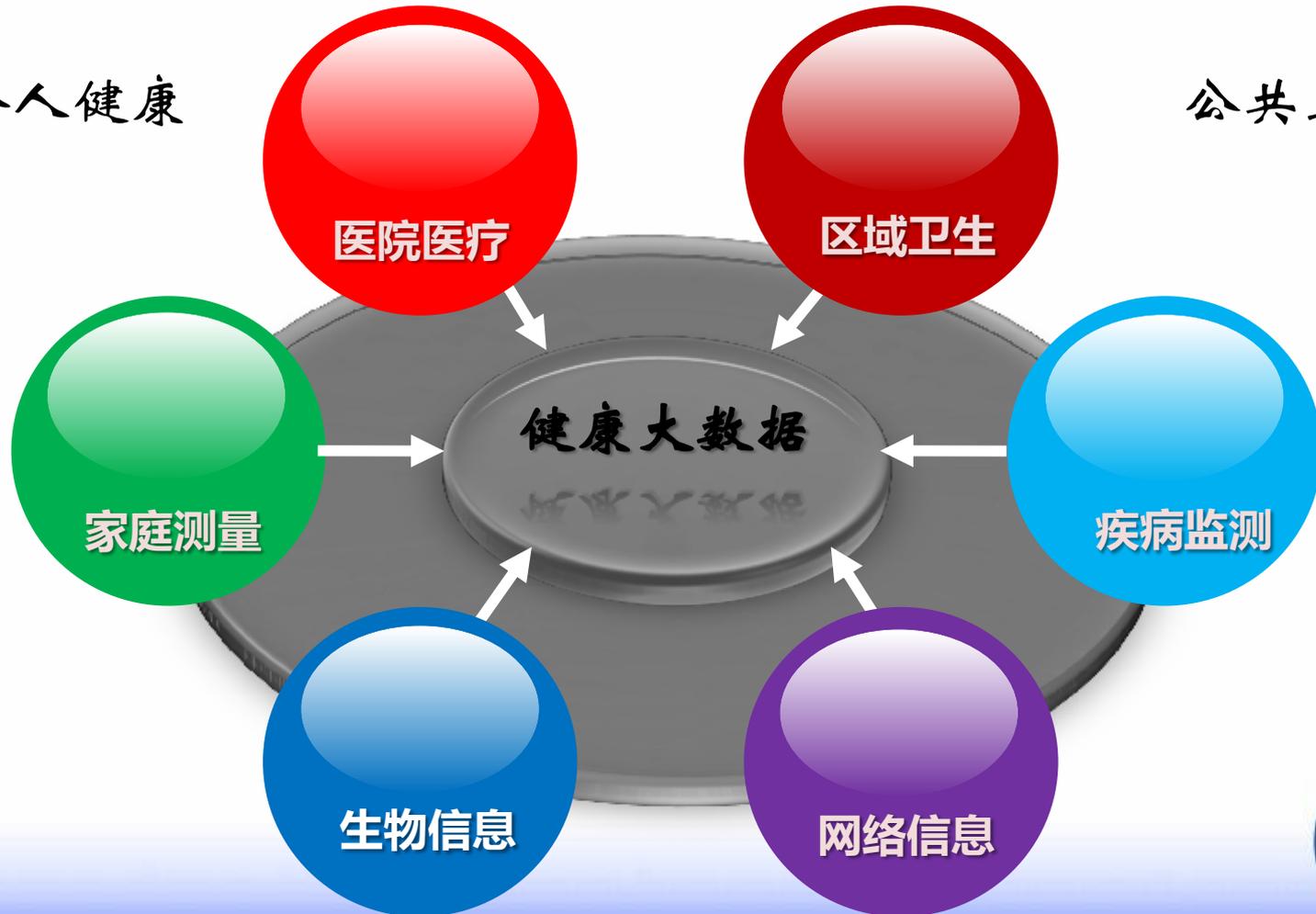


医学影像

分类

个人健康

公共卫生



区域卫生服务大数据

来源

- 区域协同数据
- 各级医院
- 医疗机构
- 社区卫生机构

特点

- 充分论证
- 科学规划
- 数据规范
- 数据量大

价值

- 汇集整合区域内医疗健康数据
- 流行性疾病监控
- 流行性疾病预防



疾病监测大数据

来源

- 脑卒中筛查与防治
- 环境与遗传对冠心病和脑卒中影响
- 营养和健康调查
- 出生缺陷监测研究
- 传染病及肿瘤报告

特点

- 充分论证
- 科学规划
- 数据规范

价值

- 流行性疾病预防
- 调查统计和报告



网络信息大数据

来源

- 网络信息
- 搜索信息
- 购物信息
- 消费信息

特点

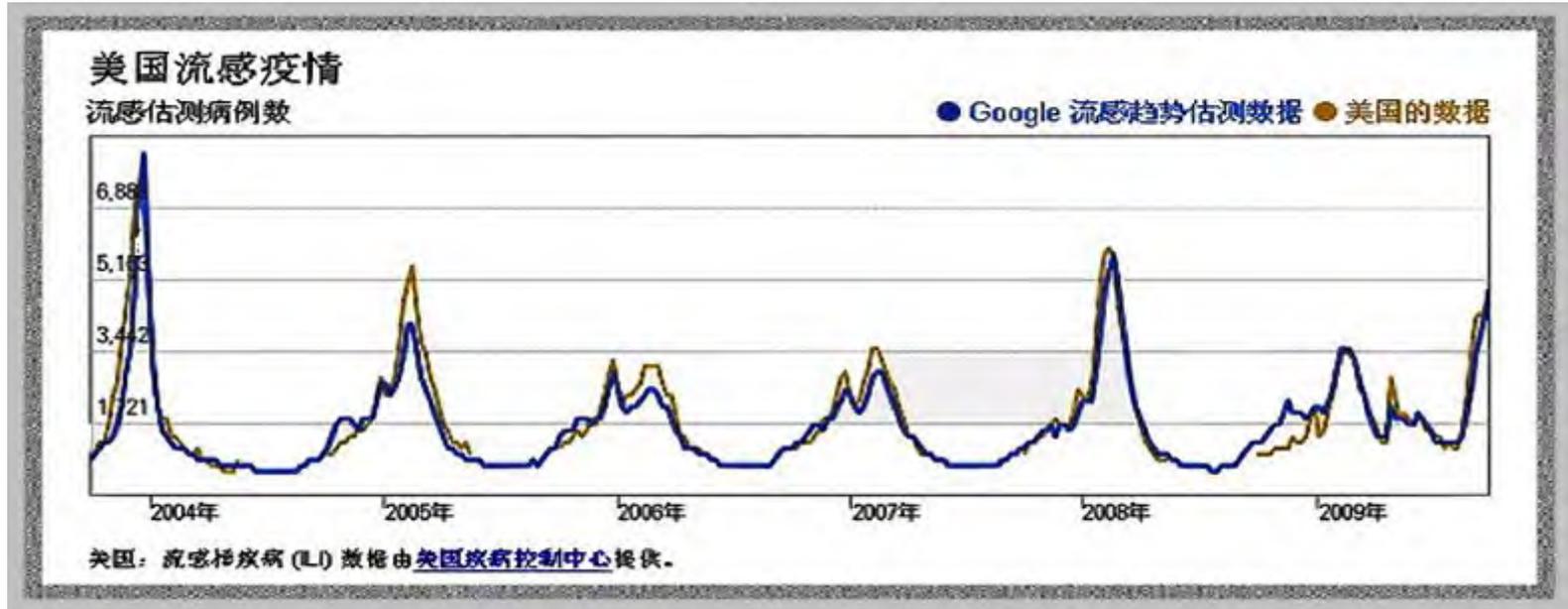
- 数据不规范
- 准确性差
- 数据量大

价值

- 流行性疾病预测
- 流行性疾病预防
- 疾病统计



谷歌流感趋势预测



“谷歌流感趋势”利用几十亿条网上的搜索查询内容，能够准确地判断流感的扩散区域。每日更新流感趋势报告，为流感的爆发提供早期预警



医院医疗大数据

来源

- 门急诊记录
- 住院记录
- 影像记录
- 实验室记录
- 用药记录
- 手术记录
- 随访记录

特点

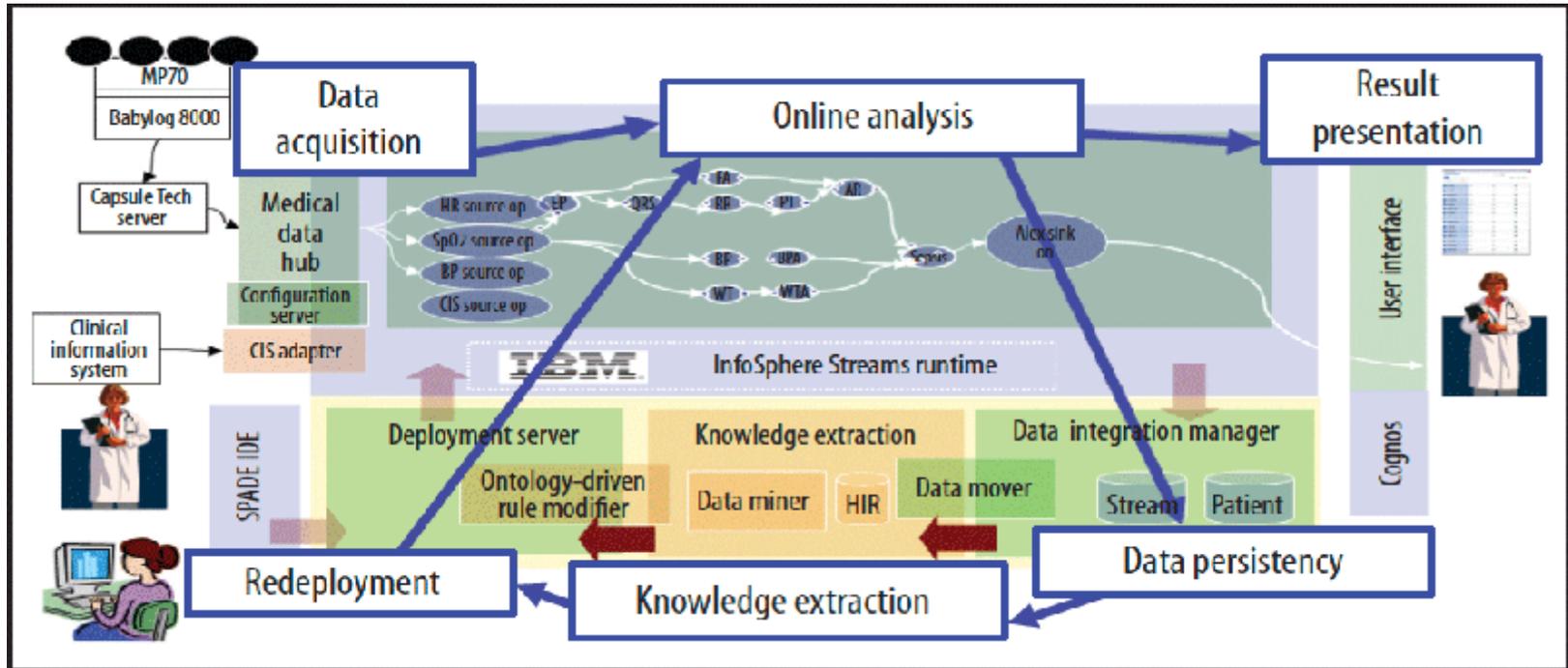
- 大多数用医学专业方式记录下来
- 数据规范和可靠
- 临床实践自然随机形式存在
- 最原始的临床记录

价值

- 就医过程的真实记录
- 临床医疗行为的痕迹
- 数据具有高价值
- 有待发掘和利用



新生儿感染监护



加拿大多伦多儿童医院 Artemis system
每秒采集新生儿生命体征1000次，一天采集量约为86.400.000次
(1000x60x60x24)



生物信息大数据

来源

基因测序

- 人体组织
- 细胞
- 血液
- 唾液

特点

- 专业性强
- 数据量大 EB
- 科学规划
- 数据规范

价值

- 揭示生命密码
- 识别疾病病因
- 个性化诊疗
- 精准化医疗



乔布斯

患胰腺癌，称为“癌中之王”，死亡率极高，90%一年内死亡，活过5年很少见

自身和肿瘤DNA测序，每隔三个月对肿瘤序列重新分析，根据变化对症下药

斯坦福、哈佛、麻省和乔布斯对胰腺基因缺陷的探究中，赢得了7年的生命



超前意识

医生发现

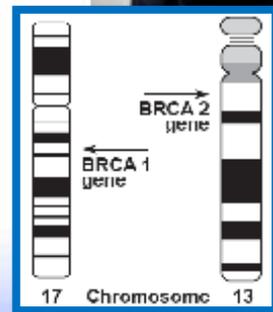
朱莉携带有“缺陷”基因BRCA1，
87%患乳腺癌，50%患卵巢癌

预防风险

不让孩子有失去妈妈恐惧，持续三
个月复杂手术，切除双侧乳腺

治疗效果

患乳腺癌的几率已下降到 **5%**



家庭测量大数据

来源

健康物联网

- 血压
- 心率
- 血糖
- 呼吸
- 睡眠
- 计步
- 运动

特点

- 连续性
- 趋势性
- 不规范

价值

- 识别疾病病因
- 防控疾病
- 个性化康复
- 全新健康管理模式

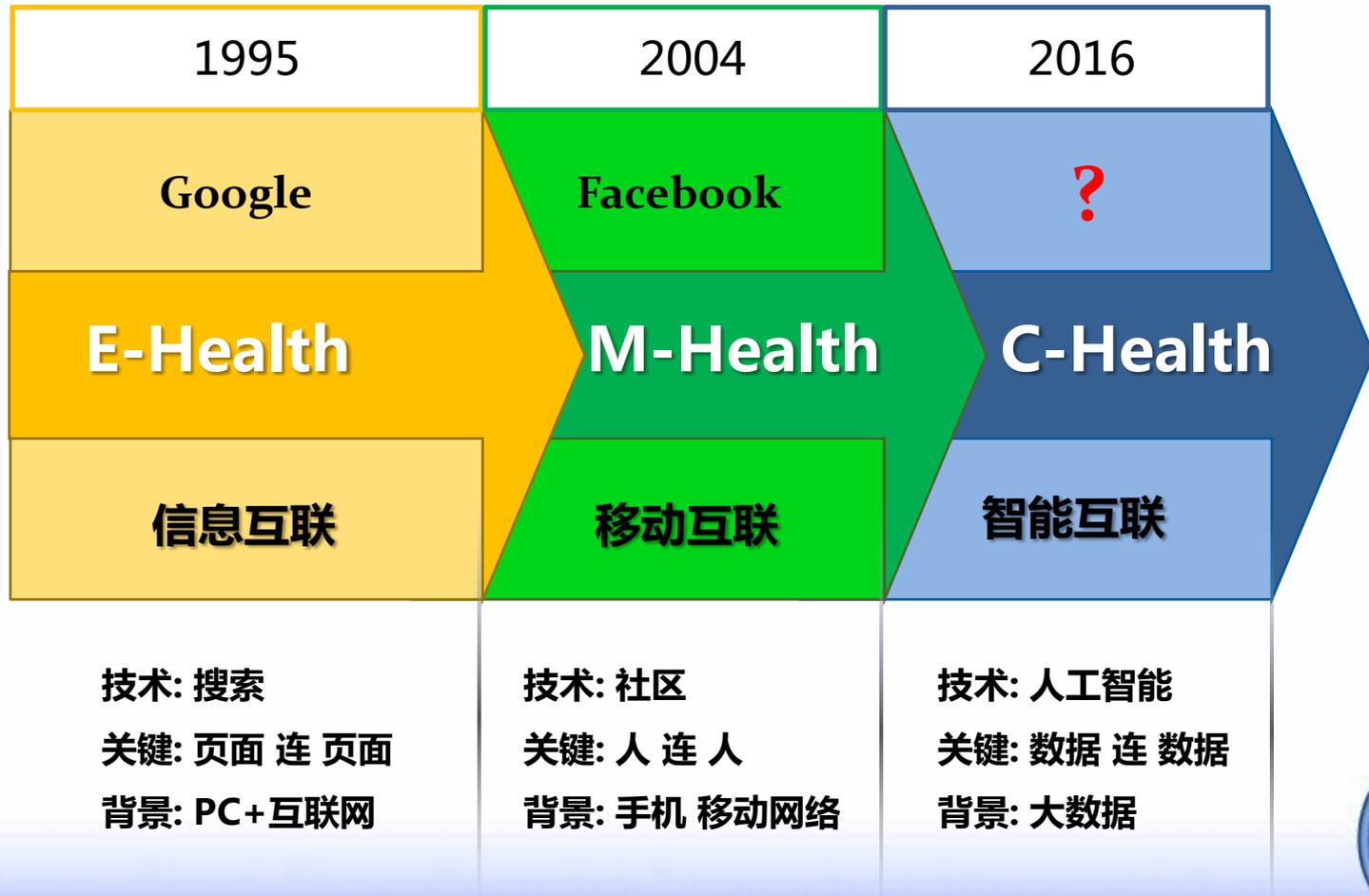


趋势

	研究载体	研究对象	研究方法
过去	专用医学设备	检测一段生理指标 简单的规则	以疾病为中心 一批数据
现在	传感器+移动终端	疾病预警 建立模型	长期监测数据
将来	可穿戴设备+云	疾病预防 培养健康行为	大数据 + 智能化 个性化 + 精准化



趋势



人工智能



智能是人类大脑的较高级活动的体现



人工智能系统(artificial Intelligence system)
具有人类智能行为的系统



自动地获取和应用知识能力、思维与推理能力、问题求解能力和自学习能力



智能健康



三阶段

被动
采集

关系型数据库，结构化数据
医疗机构采集

主动
原创

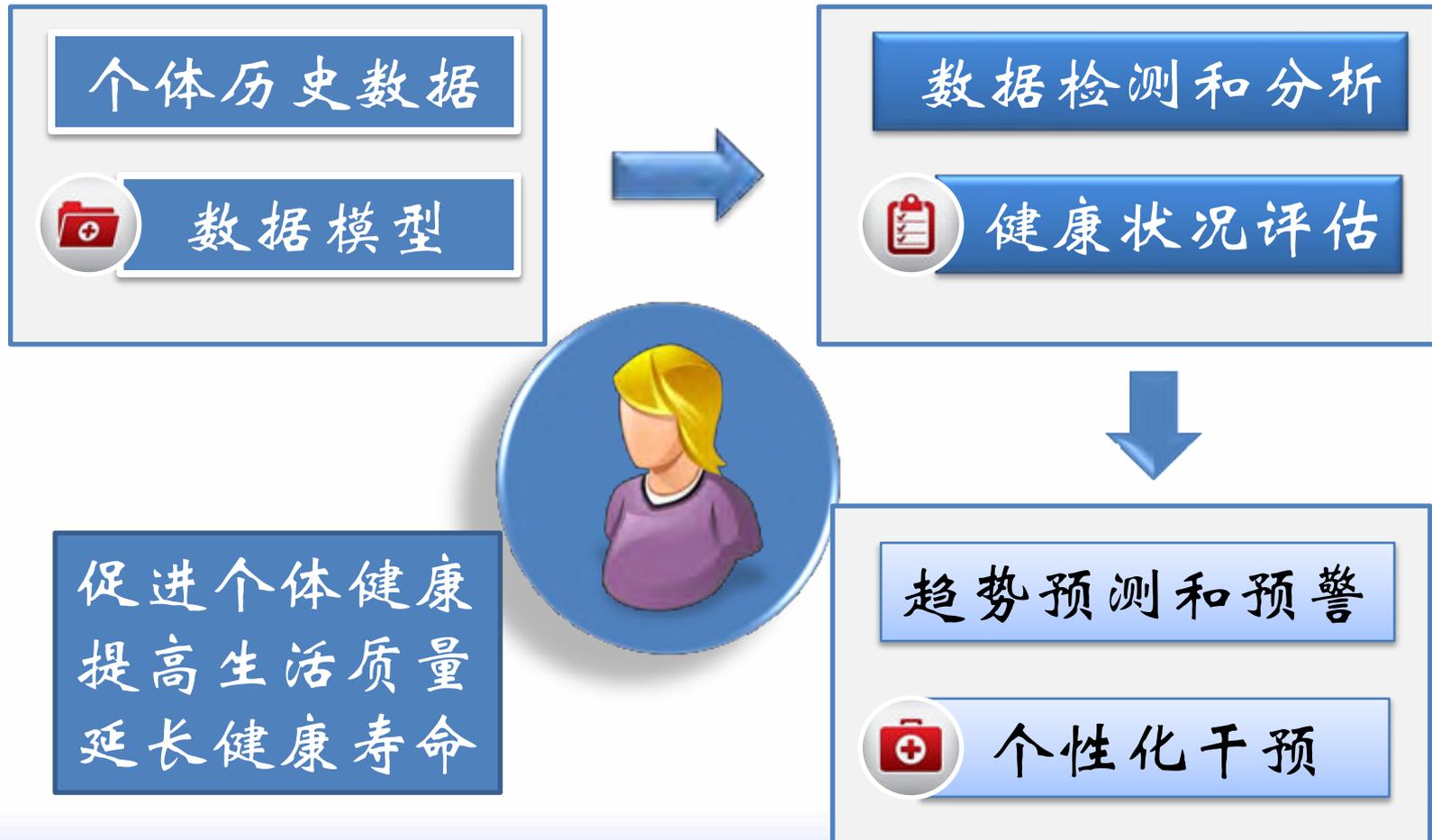
非结构化数据，个人原创
全天候网上询医互动

自动
感知

流数据，自动感知健康
健康物联网



智能健康



特点



完整的信息

全方位管理

信息· 横向到边

管理· 纵向到底

政府

社区

家庭

个人

个人健康
档案

体检档案
.....

电子健康
档案

电子病历



健康管理

卫生服务
高度依赖信息技术

健康信息
数字化

健康评估

健康危险
因素控制

慢性病管
理，个性
化健康教
育

健康计划
指导、实
施和督促

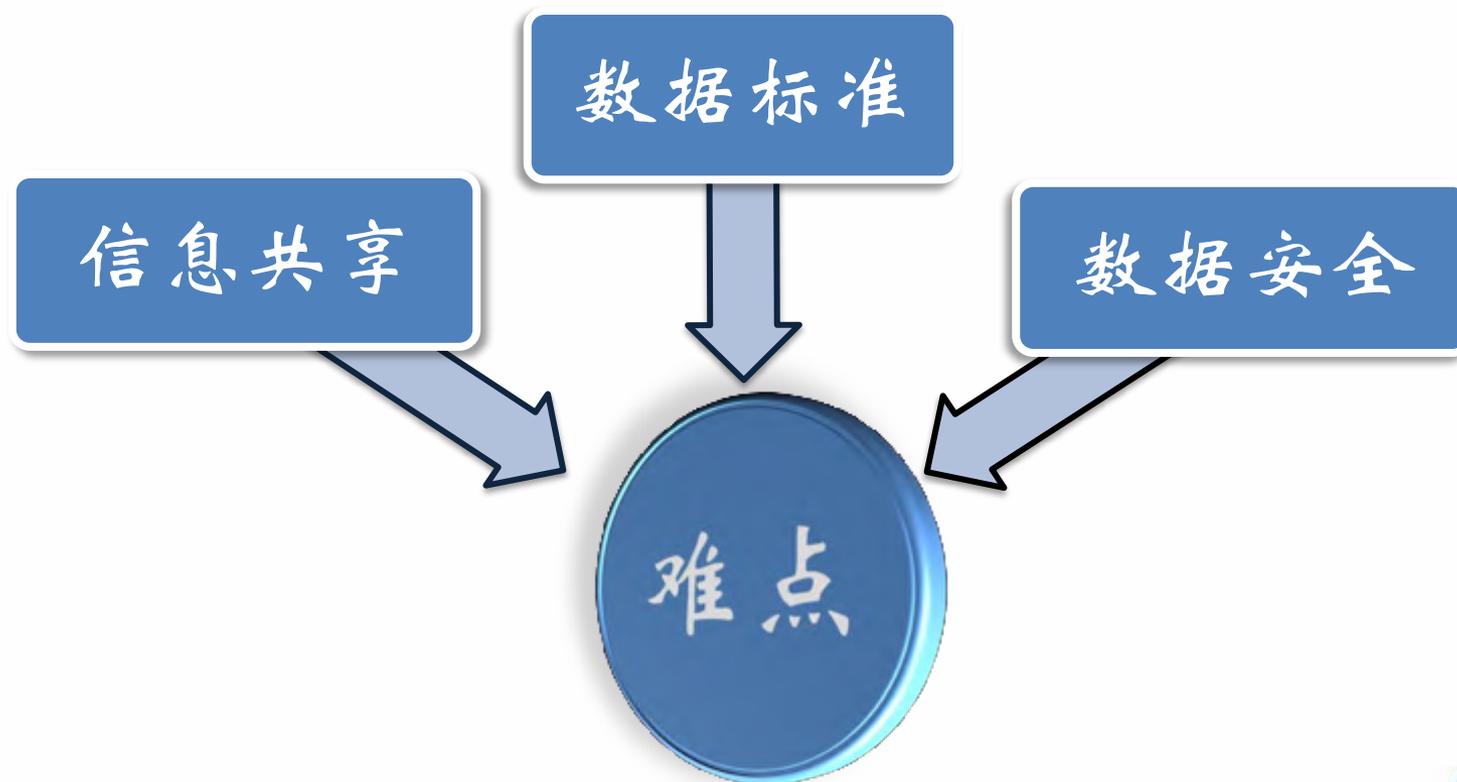
通过信息平台实现
服务市场化、规模化



O2O服务模式



难点







健康信息化

智能互联

云计算 大数据
+
人工智能

移动互联

信息互联

趋势



梦想成真

无疾而终
尽享天年

中科院上海健康科学研究所

姚志洪