



# 临床研究中的数据科学实践与探索

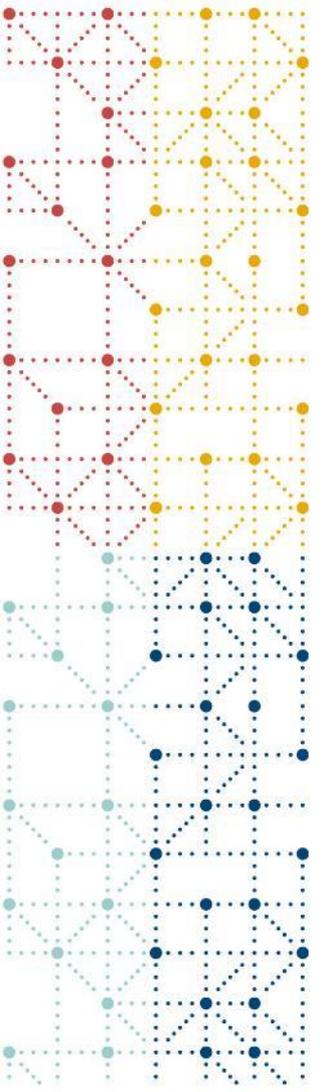
徐海涛(Raymone)

泰格医药数据科学部



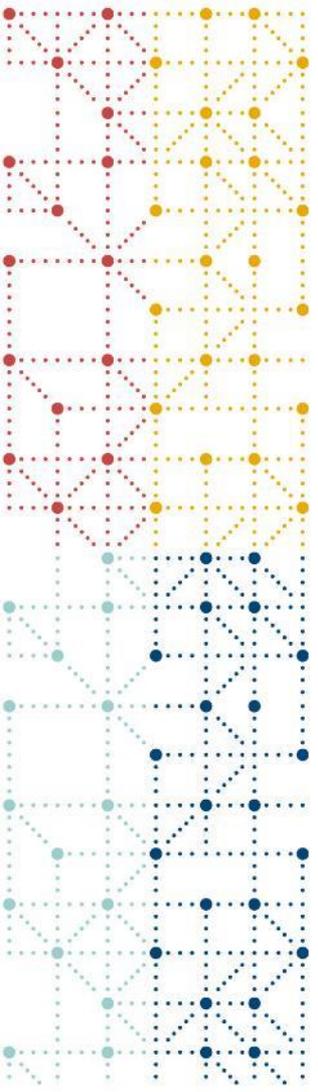
# Disclaimer and Disclosures

- *The views and opinions expressed in this presentation are those of the author(s) and do not necessarily reflect the official policy or position of CDISC.*
- *{Please disclose any financial relationship or conflict of interest relevant to this presentation here OR}*
- *The author(s) have no real or apparent conflicts of interest to report.*



# 目录

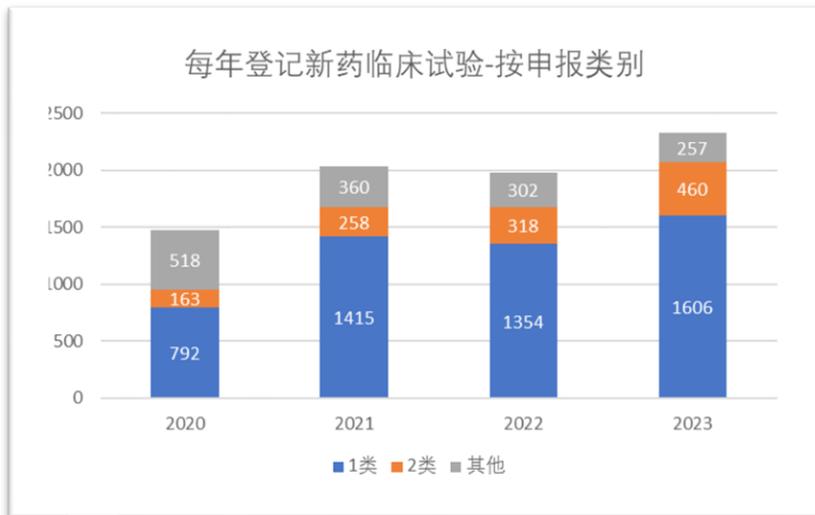
1. 临床研究之数据科学
2. 实践探索
3. 结语



## 临床研究应用数据科学的意义

- 快速洞察数据，更科学的决策
- 加强临床试验数据质量管理
- 专注于高价值工作内容，提升获得感

# 临床研究愈发复杂，对数据工作者的影响？



# 统计分析 vs. 数据科学

## 统计分析

SDTM、ADaM、TFL、CRT等

临床研究

SAS、R

通常需要以规范的格式呈现，  
并进行严格的审核和验证

## 数据科学

数据治理、数据分析与可视化、数  
据科学平台研发、中心化监查等

临床研究（过程）、企业数据

Java、Vue、Python、R

更加灵活和多样，侧重于提  
供深入的见解和预测



主要职责



范围



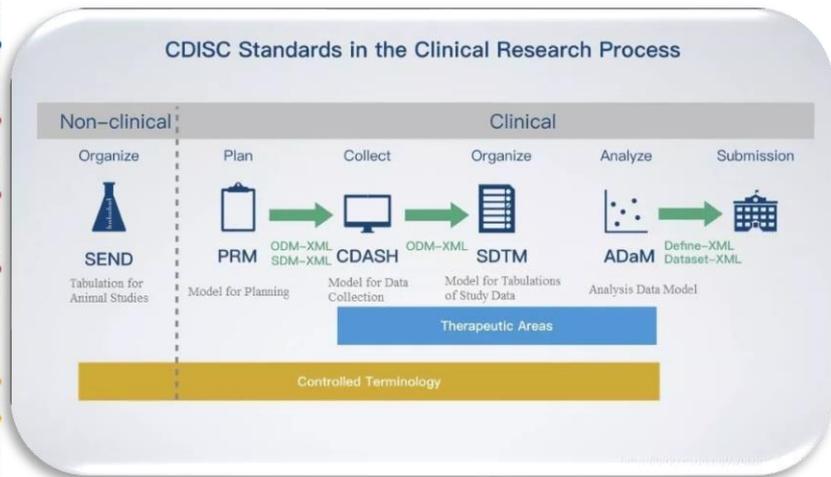
编程语言



结果呈现

# 单一的临床研究&企业系统数据流

## 临床研究数据流

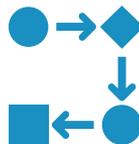


## 企业系统数据流

单一系统



数据分析



可视化看板



摘菜



洗菜、做菜



吃菜



# 复杂的企业数据高效流动、处理与应用

## 菜园子



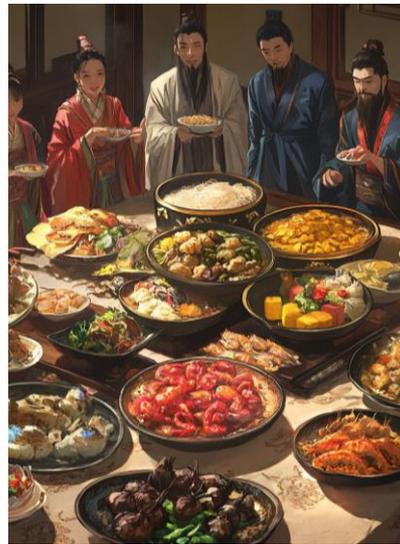
## 菜场子



## 菜厨子

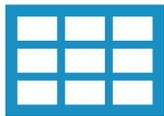


## 菜肴子



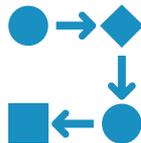
**数据源**

业务数据  
员工账号  
操作日志  
...



**数据仓库**

ODS贴源层  
DWD数据明细层  
DWM轻度汇总层



**数据分析**

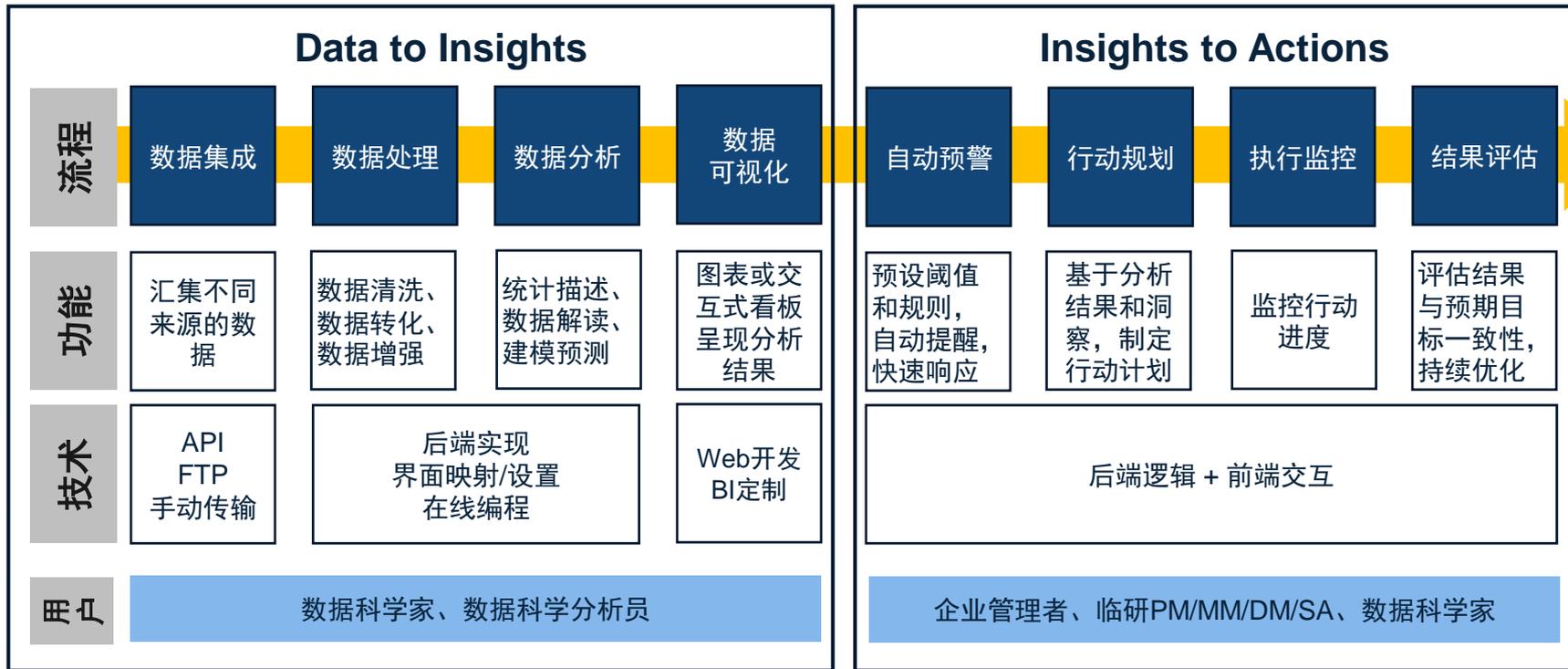
DWS数据服务层

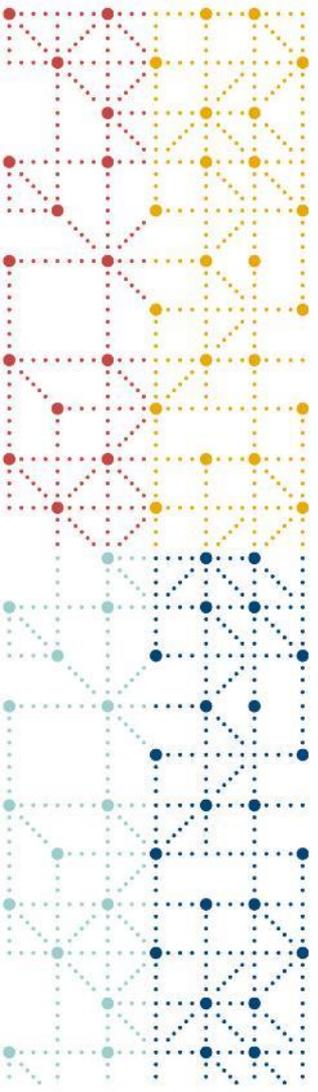


**数据应用**

实时数据看板  
数据查询产品

# Data Science = D2I + I2A

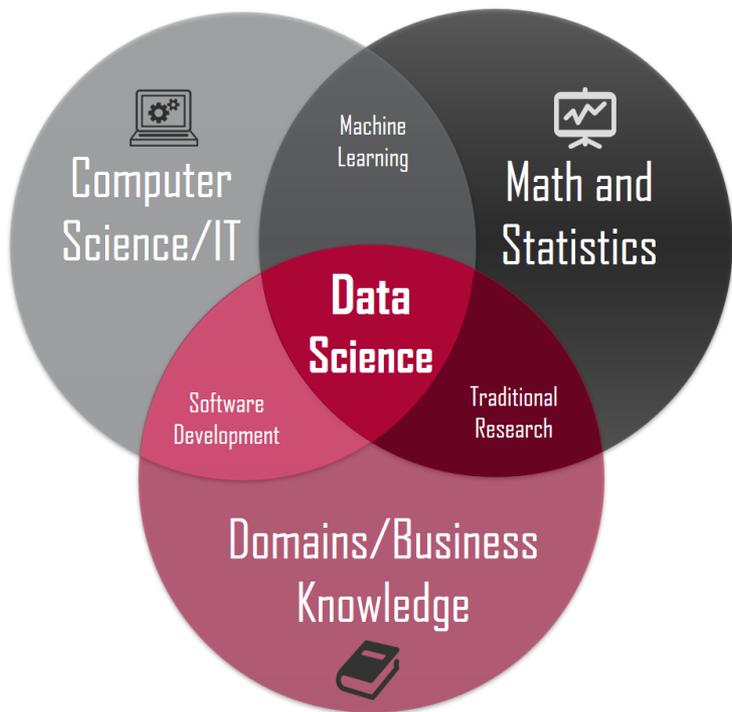




## 实践探索

- 基于风险的质量管理（RBQM）
- 中心化监查
- 医学审阅可视化

# 数据科学=业务理解+数据分析+软件技术



业务领导、业务专家、业务用户  
业务开发负责人



产品经理

软件开发（架构、前端、后端）

测试、验证



数据科学家

数据分析师

BI工程师

# RBQM - 研究中心的风险预警与评估

AD\_SRM\_DEMO / Risk Review / Site Risk Monitoring Current Project Role: CMI

Study: Site Batch Analysis Batch #8 (Cut-off Date: 2024-06-03)

Site	Region	Subject	Site Risk Score	Functional Comments	Adaptive Decisions	SDV Comp	Low Critical	Critical Data	High Query	Unmanv Query	Out of Visit	High Impair	High Nonim	High SAE	High AE Rate	Low AE Rate	High Dropt	High Solut	Low Er
5311 Hospital General Universitario Gregorio Mara...	ARG	2	37	2	Adjust On...	●	●	●	●	●	●	●	●	●	●	●	●	●	●
2026 Dana-Farber Cancer Institute	USA	1	34	1	None	●	●	●	●	●	●	●	●	●	●	●	●	●	●
2080 Sanford Roger Maris Cancer Center	USA	1	34 ↑	1	Adjust On...	●	●	●	●	●	●	●	●	●	●	●	●	●	●
2014 Ocala Oncology Center	USA	1	34	2	None	●	●	●	●	●	●	●	●	●	●	●	●	●	●
2081 BRCR Tamarac	USA	1	32 ↑	1	Adjust On...	●	●	●	●	●	●	●	●	●	●	●	●	●	●
5304 HOSPITAL DE VALME	ARG	1	32	1	Adjust On...	●	●	●	●	●	●	●	●	●	●	●	●	●	●
2002 UCLA Hem/Onc- Clinical Research Unit	USA	6	31	1	Adjust Next...	●	●	●	●	●	●	●	●	●	●	●	●	●	●
5306 Hospital regional universitario malaga	ARG	1	31	0	Adjust On...	●	●	●	●	●	●	●	●	●	●	●	●	●	●
2104 Cross Cancer Institute	CAN	2	30	2	Adjust On...	●	●	●	●	●	●	●	●	●	●	●	●	●	●
5313 Hospital Teresa Herrera (C.H.U.A.C)	ABW	2	30 ↑	2	None	●	●	●	●	●	●	●	●	●	●	●	●	●	●

\* 源自泰格RBQM系统

- ✓ 关键风险指标提示重要待办，优先关注影响质量的关键数据与流程
- ✓ 研究中心风险分级（高风险、中风险、低风险），基于研究中心风险等级采取风险相称的质量管理（/监查）

# 多中心临床研究的监查职责分配与一体化



## 数据管理审核

检查数据的有效性、一致性、缺失和正常值范围等，并通过数据质疑进行数据清洗。



## 统计审阅

评估可能影响统计结论有效性和可解释性的数据问题或事件，如方案偏离、患者脱落、伴发事件等



## 医学监查

从医学角度审核研究数据的科学性和一致性，关注患者安全。

- 入排审阅
- 疗效数据审阅
- 安全数据审阅等



## 集中监查

评估研究中心的质量与质量风险；识别系统性数据问题和数据可靠性问题；挑选需要有因现场监查的研究中心/流程

## 研究中心监查（/CRA监查）

通过向研究中心实施临床试验提供监督、审核和支持，确保试验过程的合规性与数据完整性。  
SDV/SDR、沟通协调...



# Clinical BI Platform

## 交互式可视化数据审阅

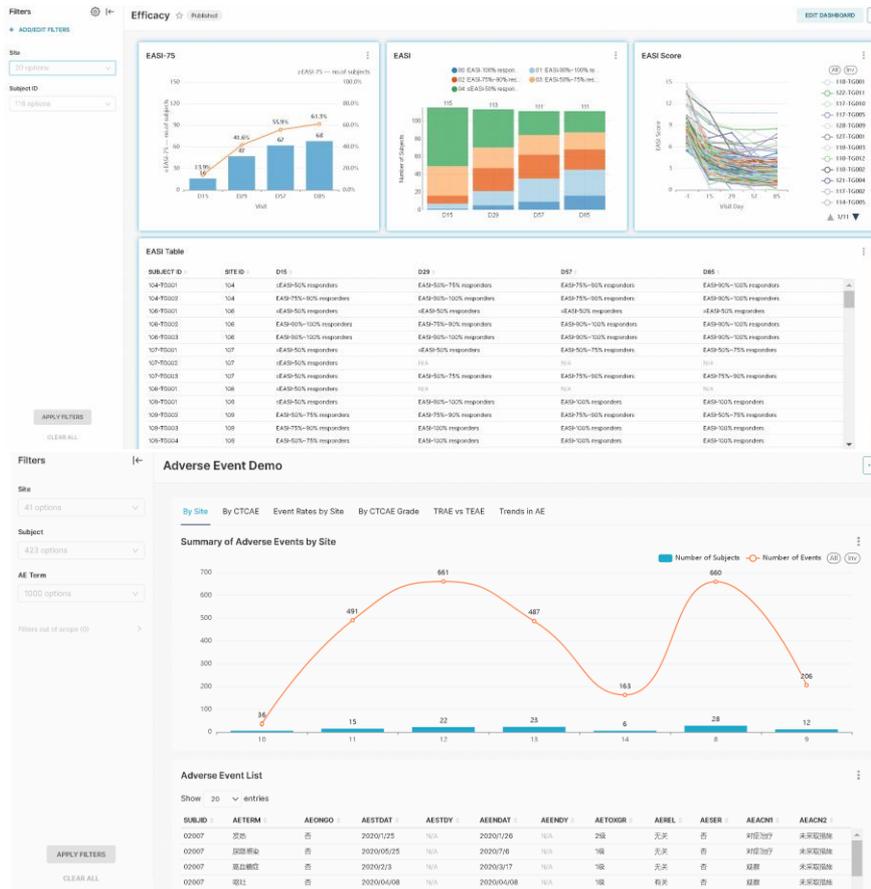
- ✓ MM-医学审阅：患者入排标准、疗效、安全性、受试者档案（Patient Profile）；
- ✓ PM-项目管理：项目进展、中心风险；
- ✓ DM-数据管理：人工数据核查、数据管理周报；
- ✓ SA-统计分析：人群分布、疗效、安全性。

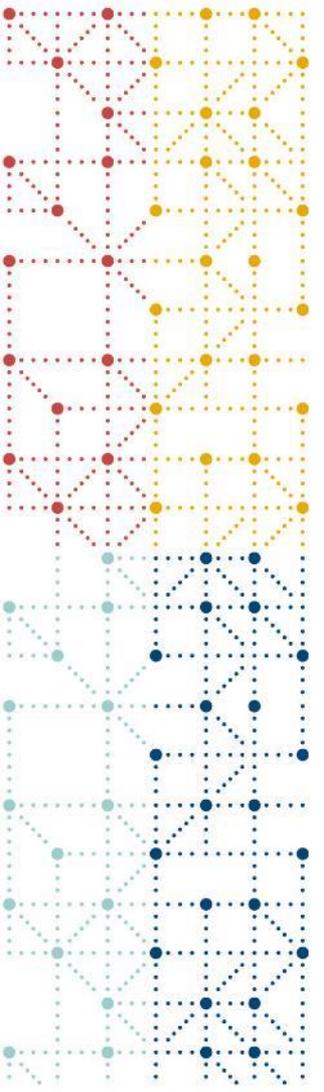
## 数据审阅+数据探索

- ✓ 试验汇总、亚总汇总、中心汇总；
- ✓ 患者呈现；
- ✓ 特别关注事件/数据；

## 自动提醒&预警

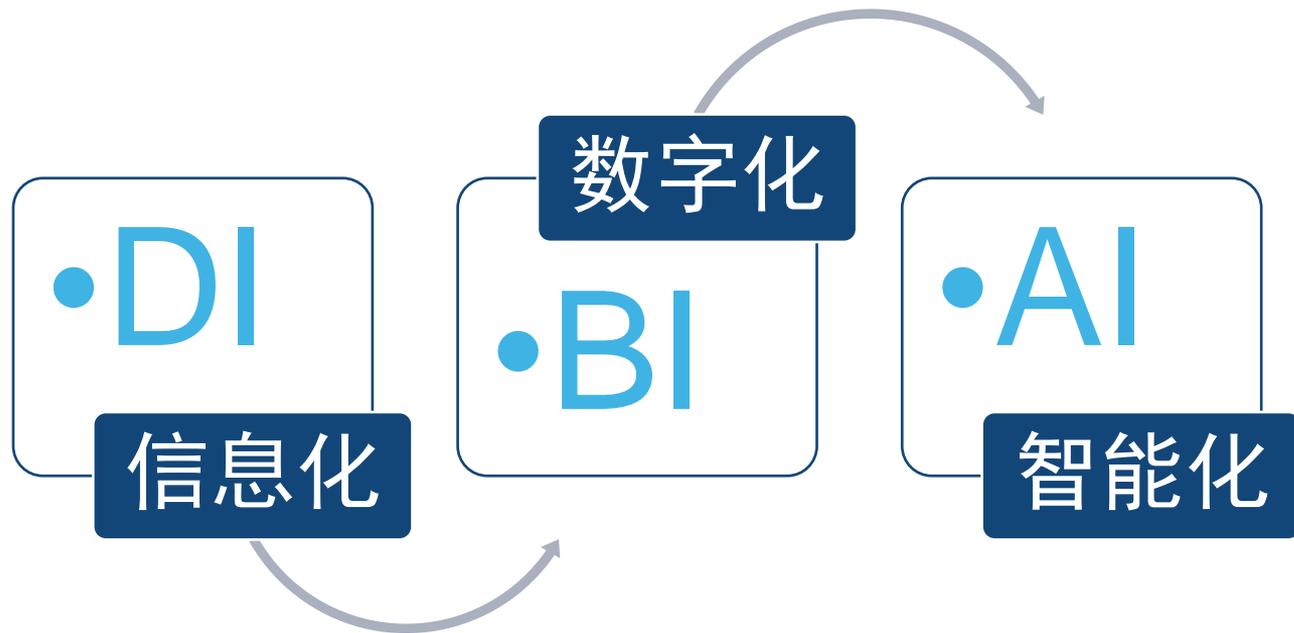
- ✓ 预设阈值或规则；
- ✓ 自动邮件推送；





# 结语

# 数据科学是推动数字化转型的关键力量



DI = Digital Information; BI = Business Intelligence; AI = Artificial Intelligence

# 数据科学是科研精神与互联网（用户）思维的激情碰撞

临床科研

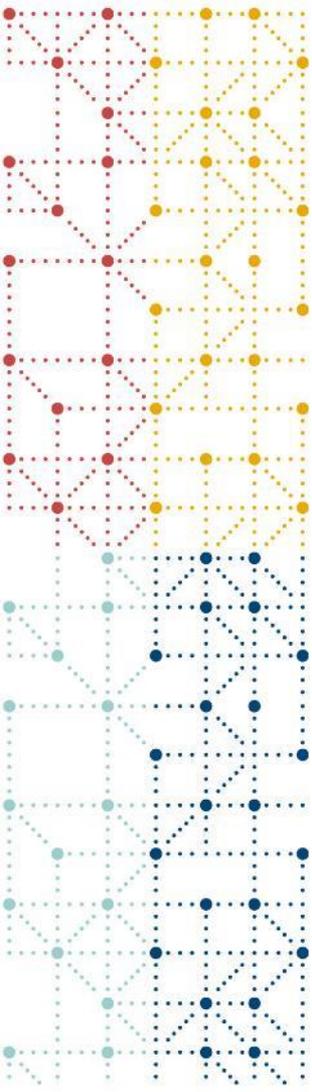


互联网



- ✓ 强调严格严谨
- ✓ 数据量较小，需要精确和高质量数据
- ✓ 业务理解

- ✓ 强调用户体验
- ✓ 数据量巨大，包括用户行为和操作日志
- ✓ 用户思维、技术与算力



感谢聆听！



cdisc