



## CDISC China update and Knowledge Frame of Clinical Data Science

Victor Wu  
北京迪时咨询有限公司  
Victor.wu@datascie.com



## Meet the Speaker

Victor Wu

**Title:** EVP, Data Science

**Organization:** 北京迪时咨询有限公司

Victor WU(吴崇胜), Ph.D., CDISC中国协调委员会(C3C)主席, CDISC官方认证SDTM及ADaM讲师; 北京迪时咨询有限公司共同创立者。在数据标准化翻译与推广、流程规范化、自动化工具开发方面有丰富的经验, 力图为行业做更多贡献。



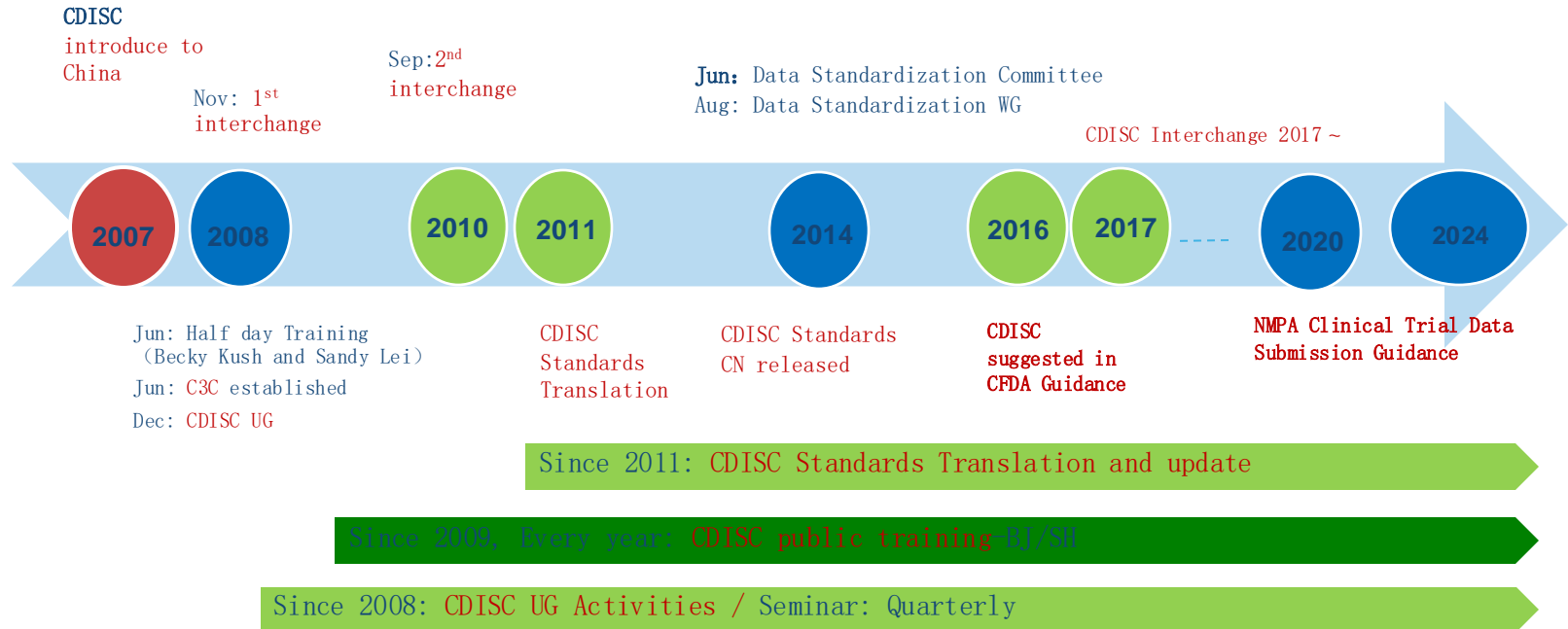
# Agenda

1. CDISC in China
2. Clinical Data Science
3. Tools / Automation / AI

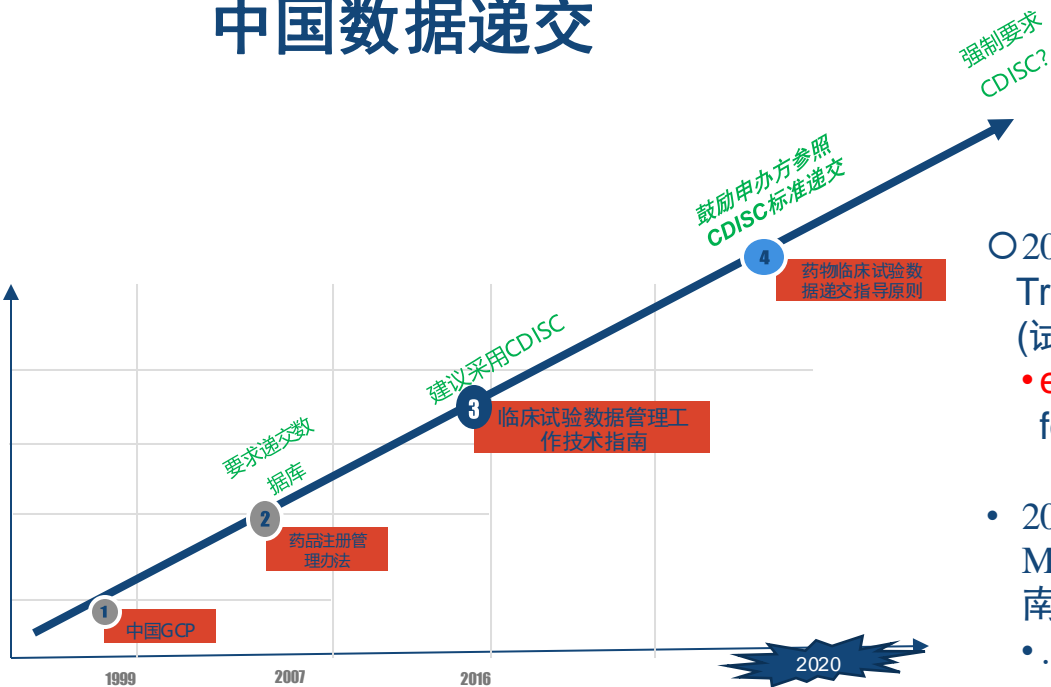


# 1. CDISC in China Update

# CDISC in China (2007-2021)



# 中国数据递交



○ 2020 Guideline on the Submission of Clinical Trial Data 《药物临床试验数据递交指导原则(试行)》

- encouraged to submit clinical trial data following CDISC standards
- 2016: Technical Guidance on Data Management 《临床试验数据管理工作技术指南》
- ...Suggested to adopt CDISC standards to submit raw database and analysis database.
- 2007: Provisions for Drug Registration 《药品注册管理办法》
- database submission was required in NDA

# 《药物临床试验数据递交指导原则（试行）》（NMPA 2020）

国家药监局药审中心关于发布《药物临床试验数据递交指导原则（试行）》的通告（2020年第16号）

发布日期：20200720

为规范药品注册申请人递交药物临床试验数据及相关资料，配合新修订的药品注册申报资料要求，提高药品审评效率，药审中心组织制定了《药物临床试验数据递交指导原则（试行）》（见附件）。根据《国家药监局综合司关于印发药品技术指导原则发布程序的通知》（药监综药管〔2020〕9号）要求，经国家药品监督管理局审核同意，现予发布。

化学药品、生物制品自2020年10月1日起实施。中药实施日期按国家药监局发布中药注册分类及申报资料要求的通告中相关规定执行。

特此通告。

国家药品监督管理局药品审评中心  
2020年7月20日

<http://www.cde.org.cn/news.do?method=viewInfoCommon&id=7a43c3abfde95950>



# NMPA递交准备Tips

## • 相关文件齐全、格式正确

- 数据说明
  - xml: define.xml(推荐)
  - pdf: define.pdf (不推荐)
- aCRF: pdf
  - ✘ 不是EDC数据集的aCRF
  - 而是递交的xpt原始数据集对应的aCRF
- 数据审阅说明: pdf (原始数据库/分析数据库)

## • 数据文件

- 文件格式: XPT V5或更高版本
  - XPT v5: SAS (proc copy)
  - XPT v8 (推荐): <http://support.sas.com/kb/46/944.html>
  - ✘ 注意: SAS CPORT做的xpt是不符合要求的
- Encoding: 需要在《数据审阅说明》中说明
  - UTF8: 推荐
  - EUC-CN

## • 必备数据集

- 原始数据库: 人口学数据集 – dm.xpt
  - 命名为dm; 每人一条记录
  - 必须变量: STUDYID/USUBJID/SUBJID
- 分析数据库: 受试者水平分析数据集 – adsl.xpt
  - 命名为adsl; 每人一条记录
  - 包括: 人口学、重要的基线特征/分层因素、治疗组、预后因素、重要日期、分析人群划分等信息

## • 其他数据集:

- 必须变量: STUDYID/USUBJID
- 原始数据库
  - VISITNUM、VISIT——对应
  - 数据集命名参照CDISC: ae, cm, ds, lb, vs, qs...
- 分析数据库
  - 通常以“adxxxxxx”命名
  - 分析数据集的命名应尽量与原始数据集保持对应, 如: adcm、adae、adlb等

申办方应【在数据审阅说明中】说明所用编码（Encoding, 如utf-8、euc-cn等），以避免所递交的数据集出现乱码的情形



# 外文数据库 Foreign language database

## • What to translate 翻译内容

- ✓ 数据库：CSR报表中涉及的内容 Values referred in CSR:
  - AE/MH/CM 名称: *coded values*
  - 疗效指标
  - ...
- ✓ 数据说明文件meta: 数据集标签、变量标签；衍生过程说明；涉及疗效指标的取值或编码列表
- ✓ 数据审阅说明
- ✓ aCRF：问题描述；涉及疗效指标问题的取值或编码

C3C解读: 相关文件、数据集/变量标签、CSR中图表中涉及的数据内容应为中文。

# CDISC Standards Translation

目前可以参考的CDISC中文翻译文件有：

- SDTM metadata (IG V3.2/3.3), 待发布
- ADaM metadata (IG V1.2/1.3), 待发布
- 常用受控术语文件，已发布（**更新版准备中, call for volunteers**）
  
- SDTMIG v3.2，已发布
- ADaMIG v1.1、ADaM OCCDS v1.0, 待发布
- 中医针刺治疗领域标准用户指南 (版本1.0)，已发布
- [中医冠心病心绞痛治疗领域数据标准用户指南1.0版](#)，已发布

可以在CDISC官网，<https://www.cdisc.org/translations/chinese> 下载到以上翻译文档。

# CDISC China Activities Plan 2024

- CDISC webinar
- China Webinar: June, July
  
- CDISC China Interchange, 上海30-31 Aug2024
- CDISC User Group Activities: Oct ~ Dec 2024
  - Beijing
  - Shanghai
  - ?
  
- SDTM meta translation update
- ADaM meta translation update
- CT translation update



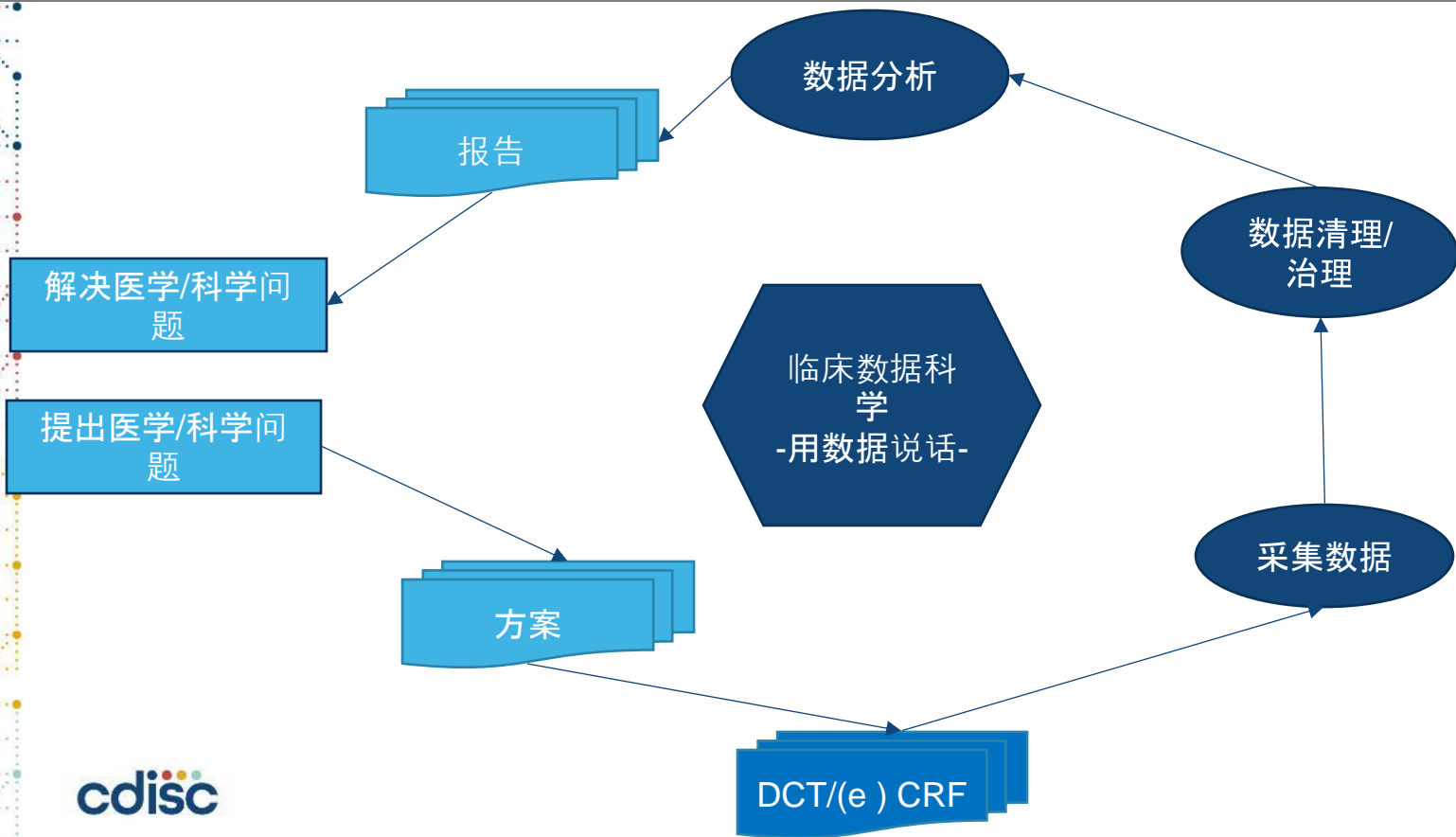


## 2. Clinical Data Science

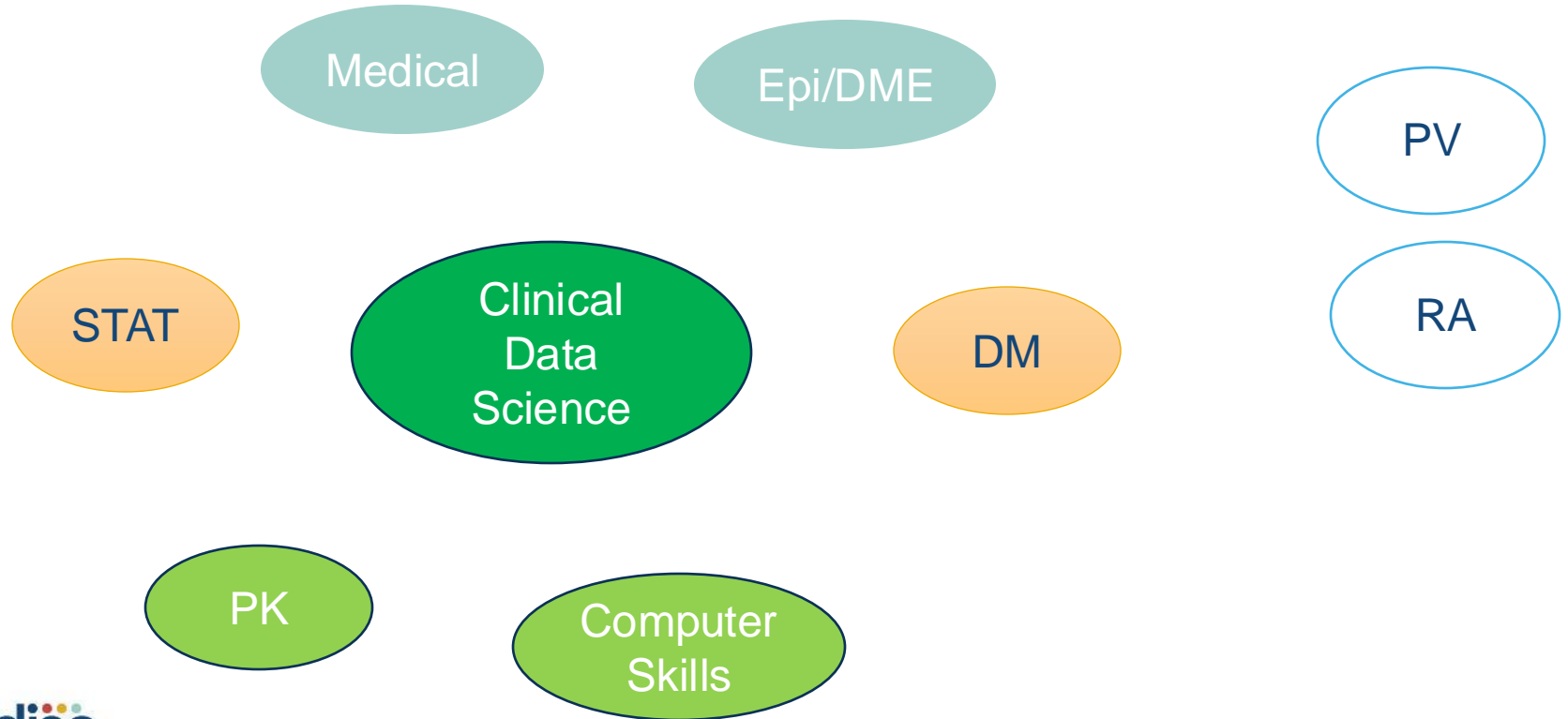


# Clinical Data Science Flow

## 临床数据科学流程



# What Clinical Data Science Covers?





## More skills are needed

- The capability to use Tools:
  - Database
  - Tools: Including AI
  - Programming





## 3. Automation / AI



# 标准化

CDISC

GCDMP

内容：标准化

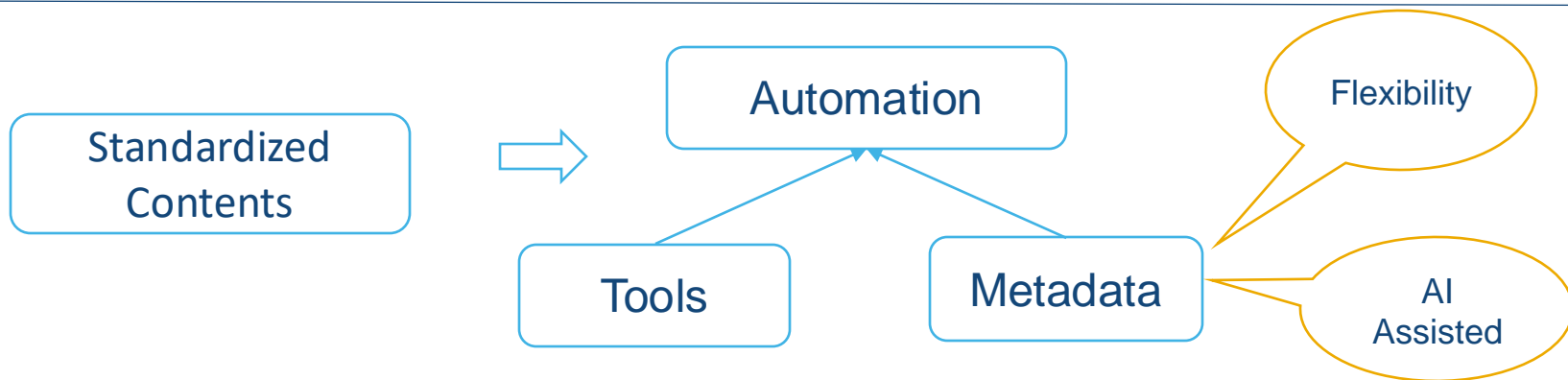
流程：规范化

交付物

过程文档

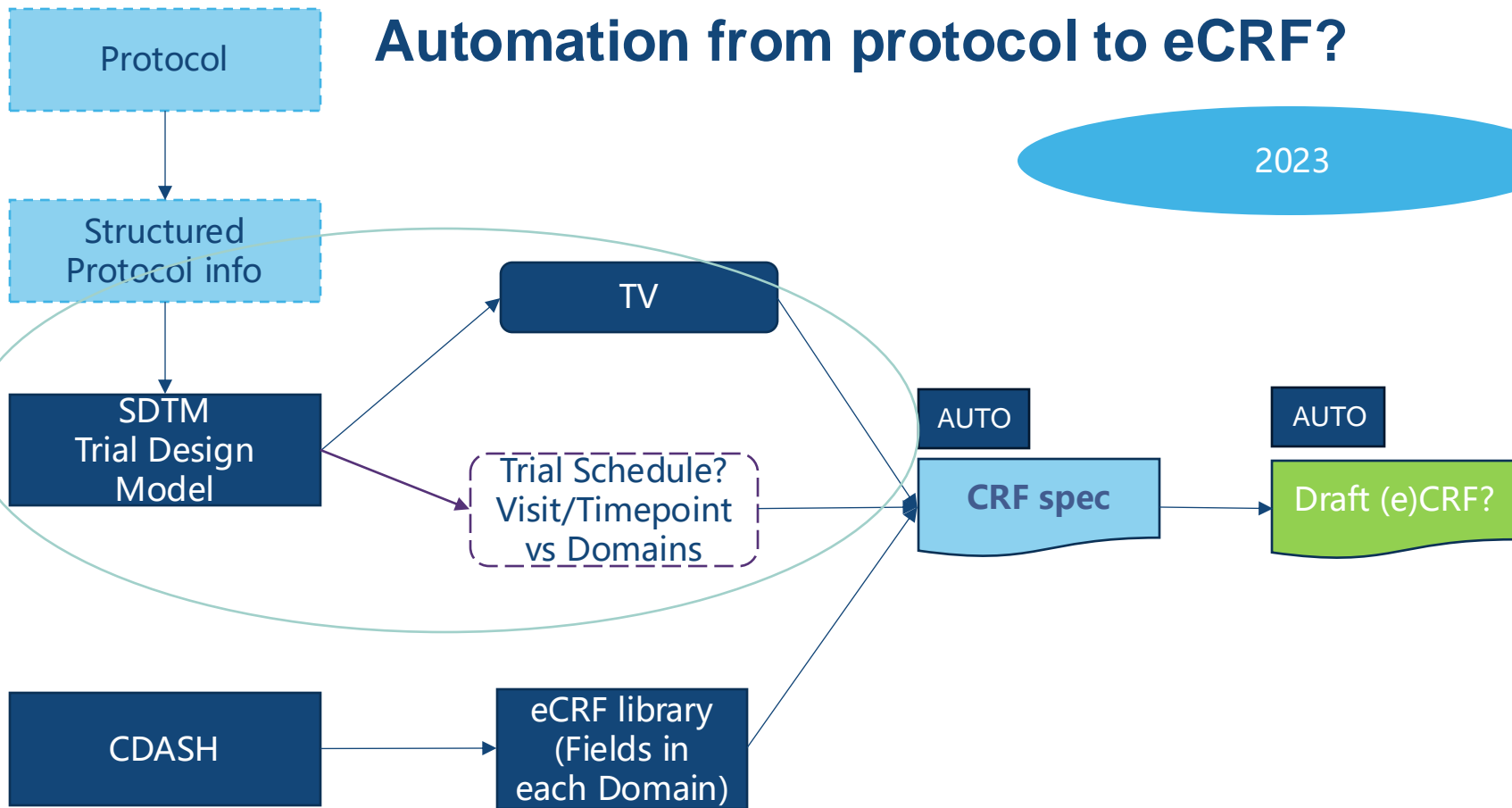
一致性

# Standards-based/Metadata-driven Automation

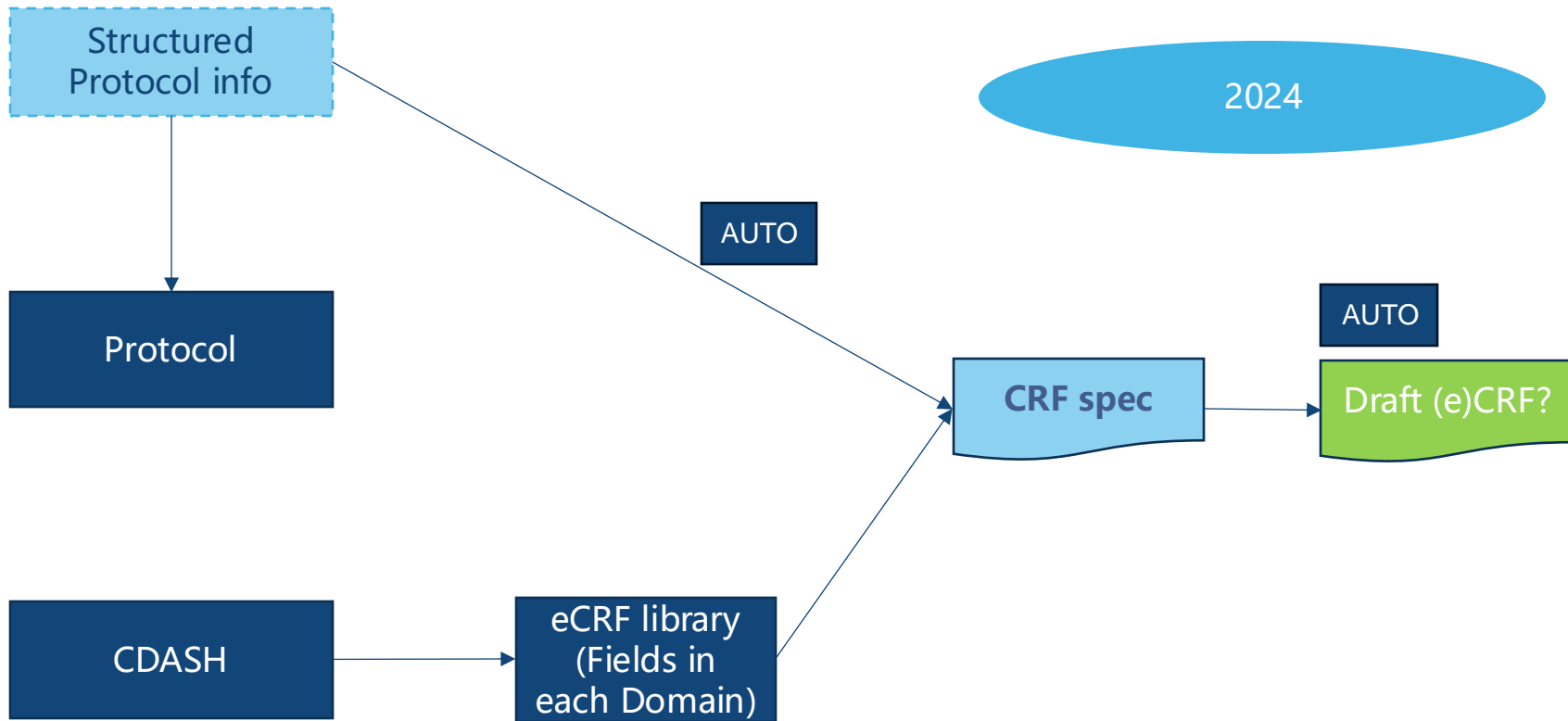


# Automation from protocol to eCRF?

2023

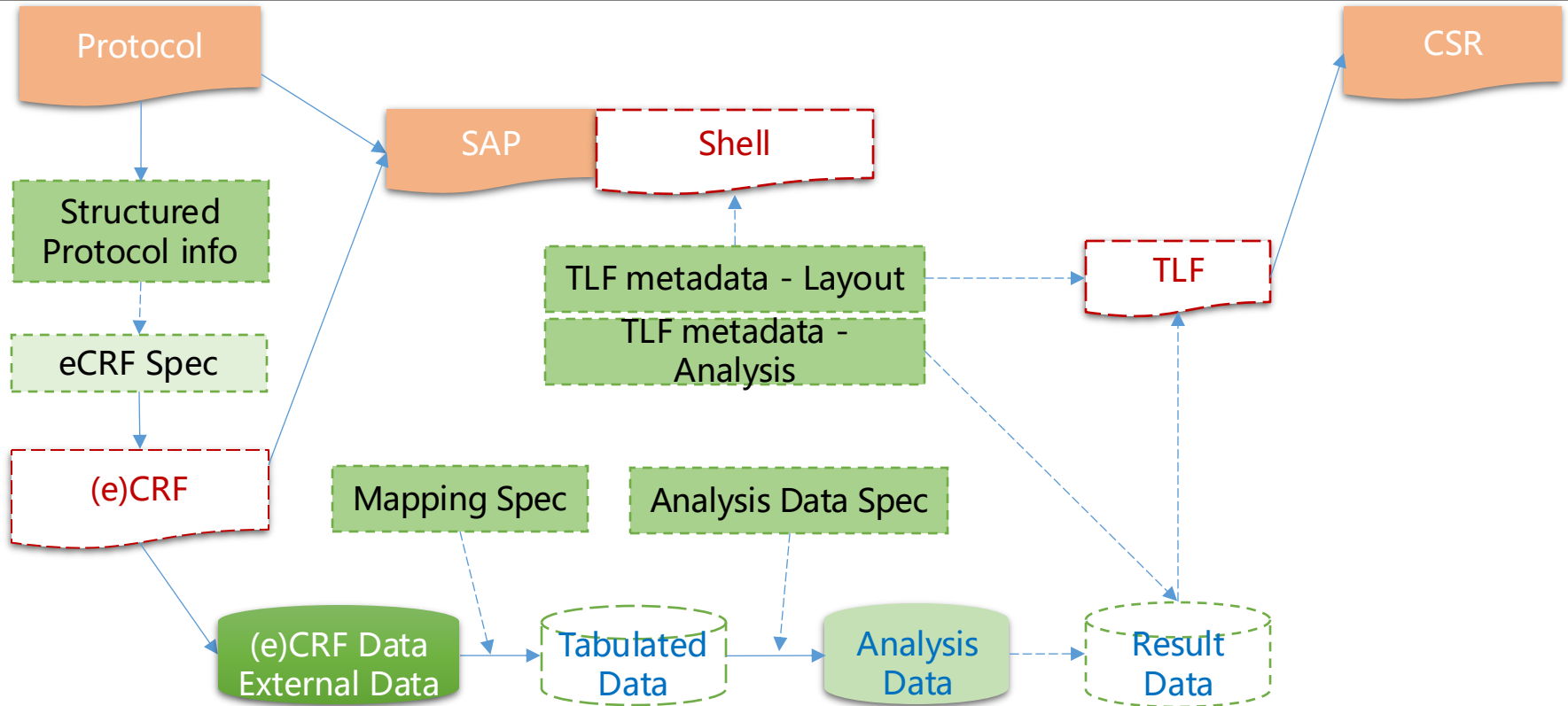


# Automation from protocol to eCRF?





# Data Flow 2023





# Data Flow 2024

Protocol

SAP

Shell

TLF

CSR

Structured  
Protocol info

eCRF Spec

Analysis Result Meta

Result Data

(e)CRF

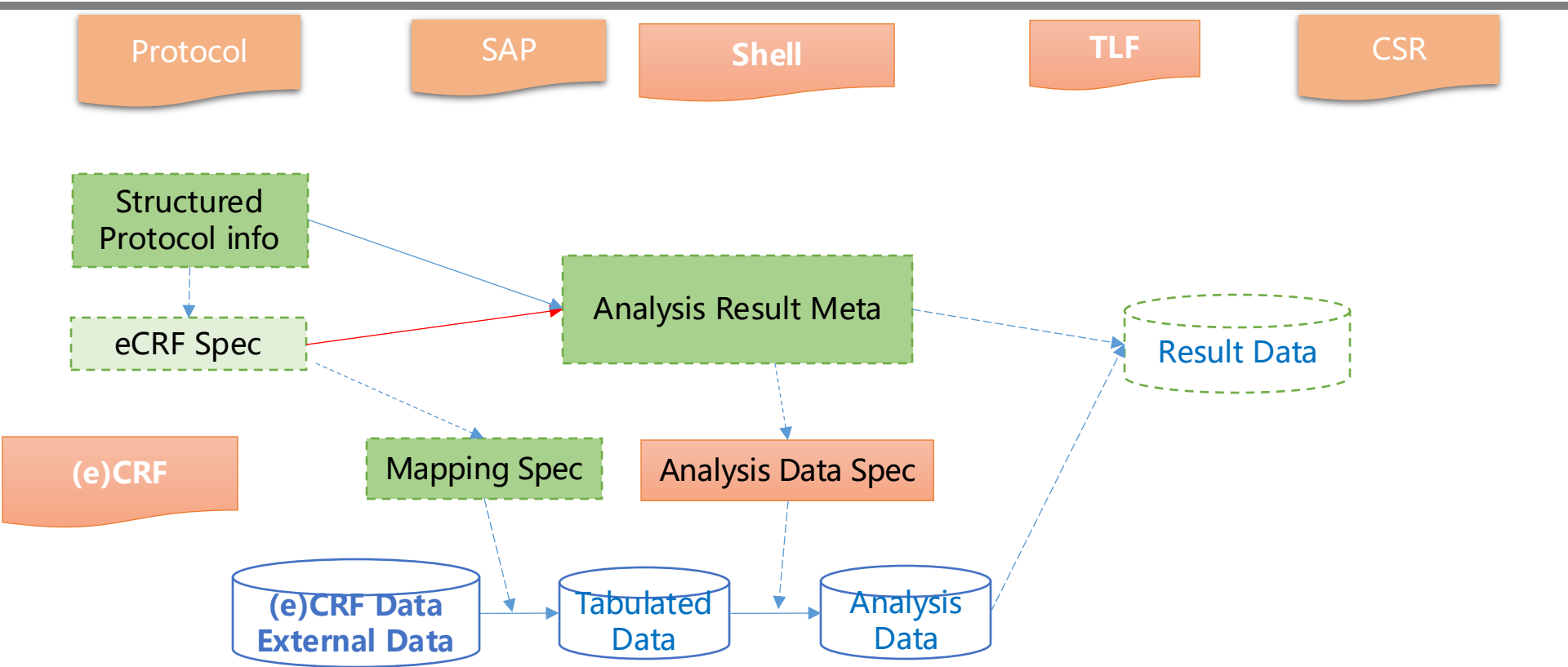
Mapping Spec

Analysis Data Spec

(e)CRF Data  
External Data

Tabulated  
Data

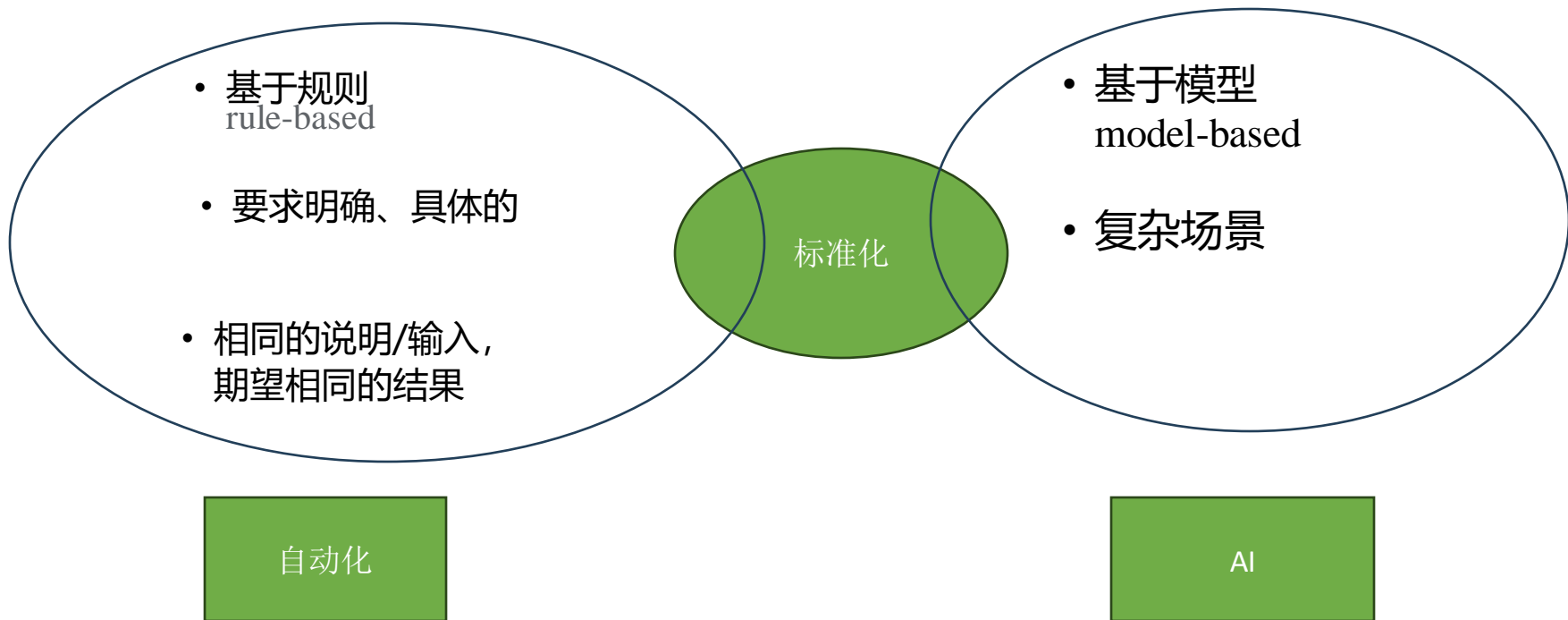
Analysis  
Data





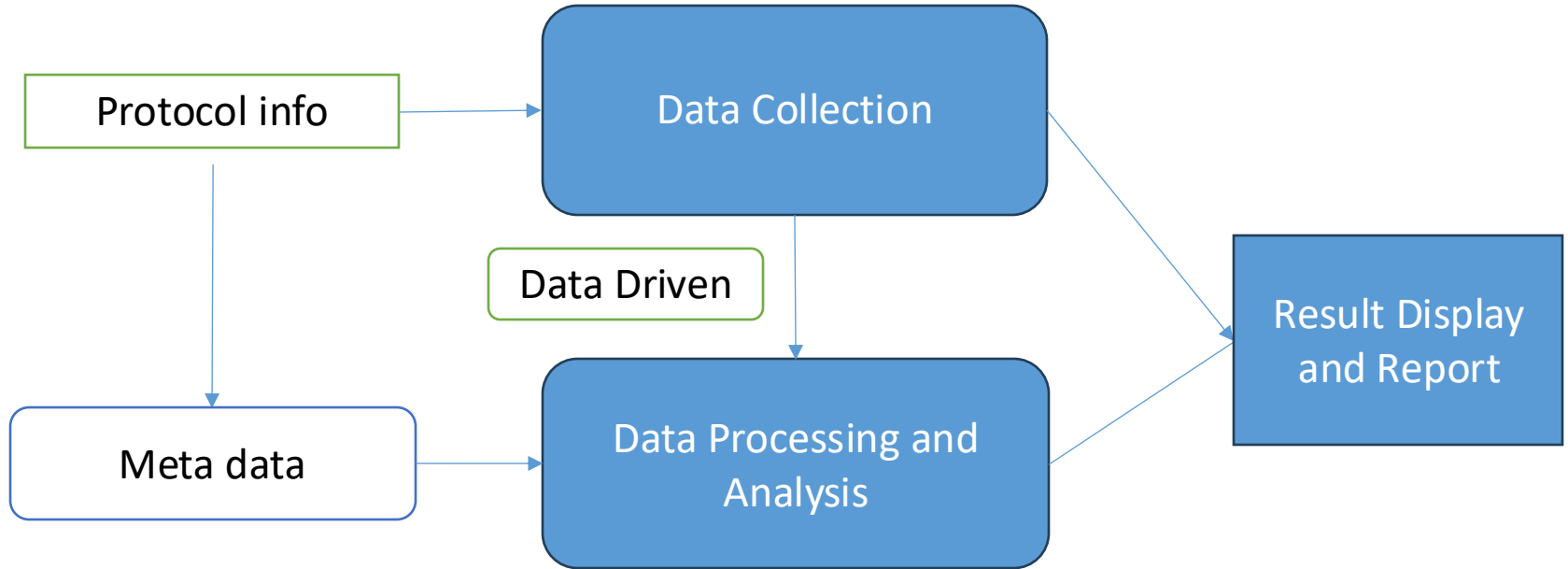


# Automation: Rule based vs Model Based



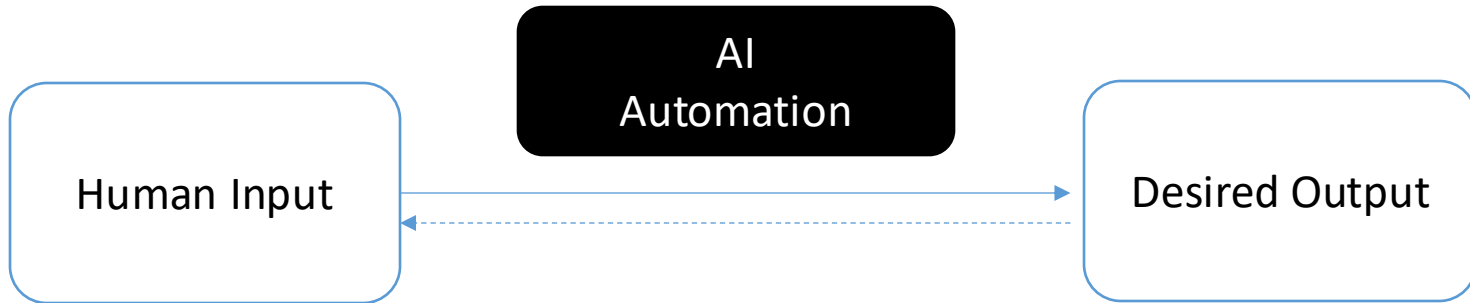
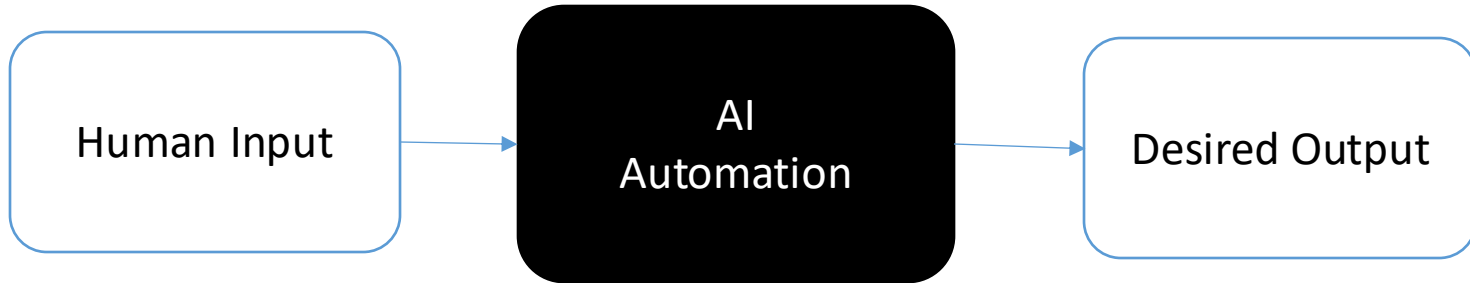


# Meta/Data Driven





# The Relationship





# What is your desired relationship?

---

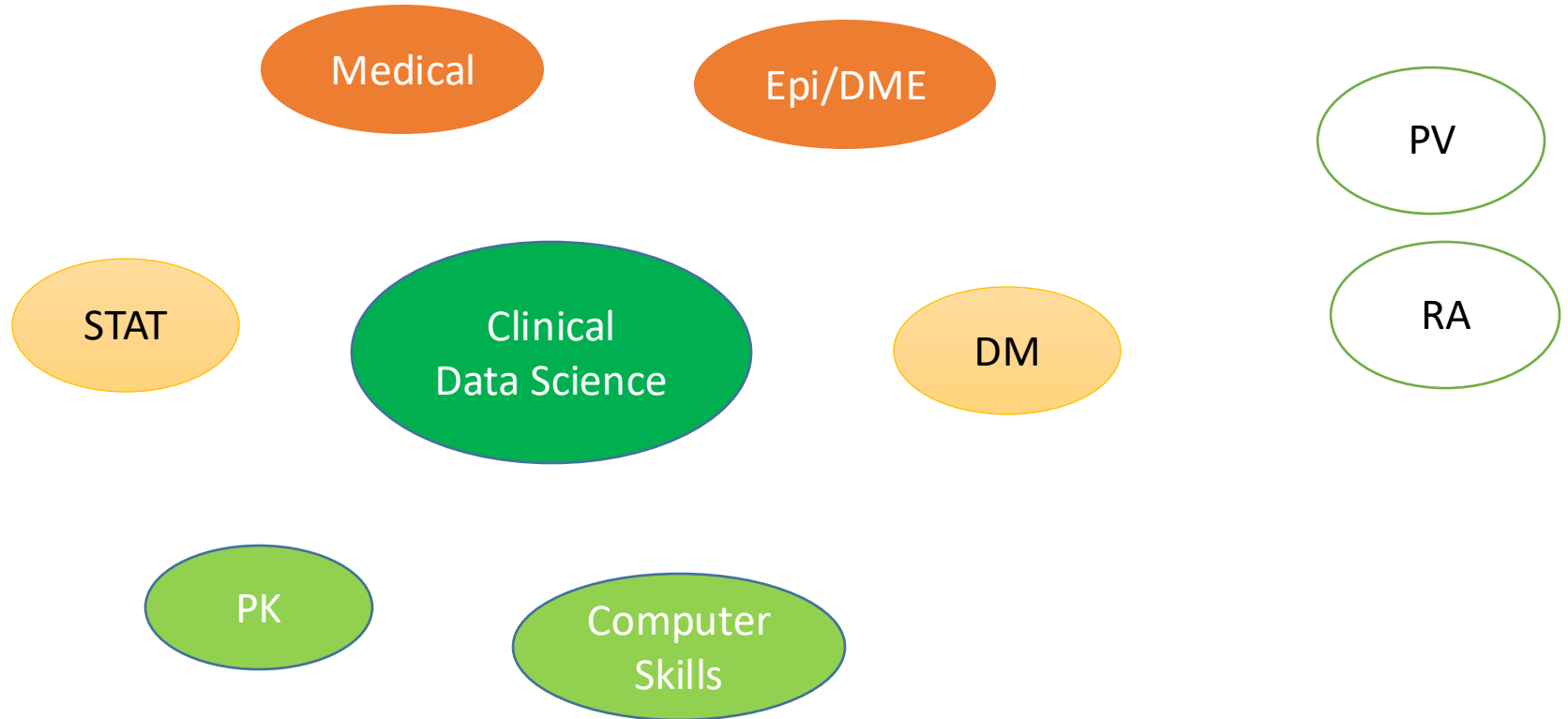
gives us  
What she can?

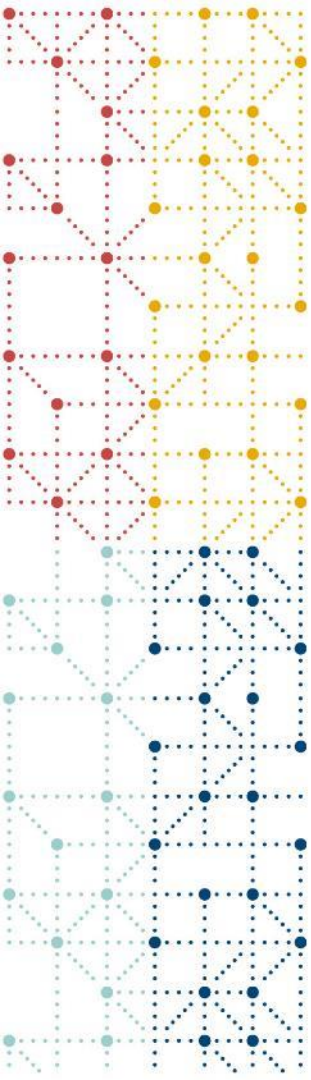
gives us  
What we need?

Soul Mate

We need to know what we  
want first

# Clinical Data Science





**Thank You!**

Victor.wu@datascie.com



**cdisc**