Technology Enabled Standards: Leveraging Technology, Open-Source Solutions, and AI to Drive Standards Automation

技术赋能标准:利用技术、开源解决方案和AI推动标准自动化

Julie Smiley, Vice President, Data Sciences, CDISC

Meet the Speaker

Julie Smiley, MS

Title: Vice President, Data Sciences

Organization: CDISC

Julie Smiley has over 25 years of experience in the pharmaceutical industry and extensive expertise with data standards and data flow automation. Throughout her career, she has led data management and database programming organizations, later shifting her focus to standards, metadata governance, and end-to-end data flow automation. As Vice President of Data Sciences at CDISC, Julie is dedicated to driving innovation in data standards, enhancing interoperability, and supporting the industry's evolving data needs. She holds a Master of Science in Computer Information Systems from Boston University, specializing in database and knowledge management.

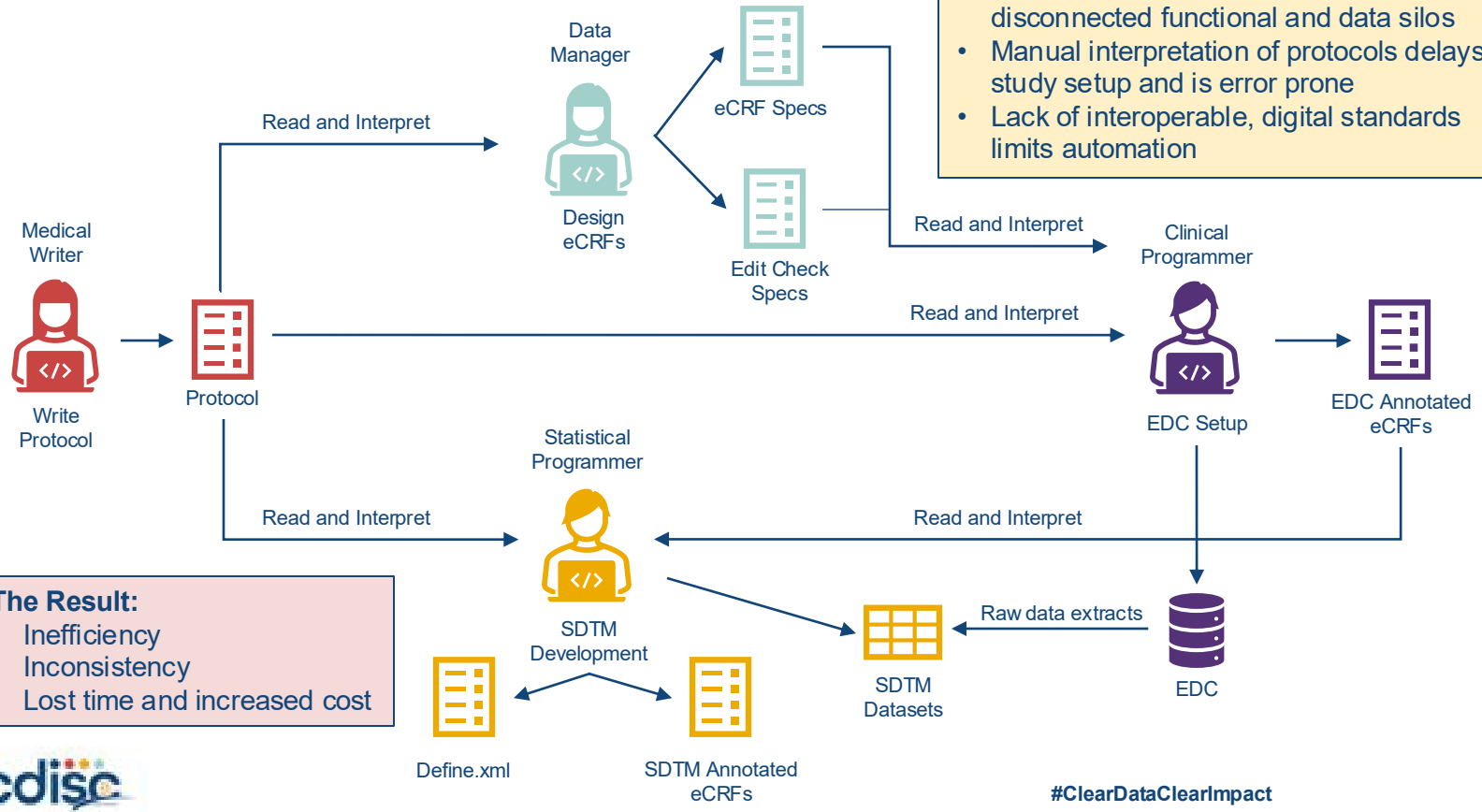
Agenda

1) Setting the Stage
2) Vision for 360i
3) 360i Technical Roadmap

Setting the Stage: The Problem

Today's Clinical Research Landscape:

- Study Design, Data Collection, Tabulations, and Analysis often operate in disconnected functional and data silos
- Manual interpretation of protocols delays study setup and is error prone
- Lack of interoperable, digital standards limits automation

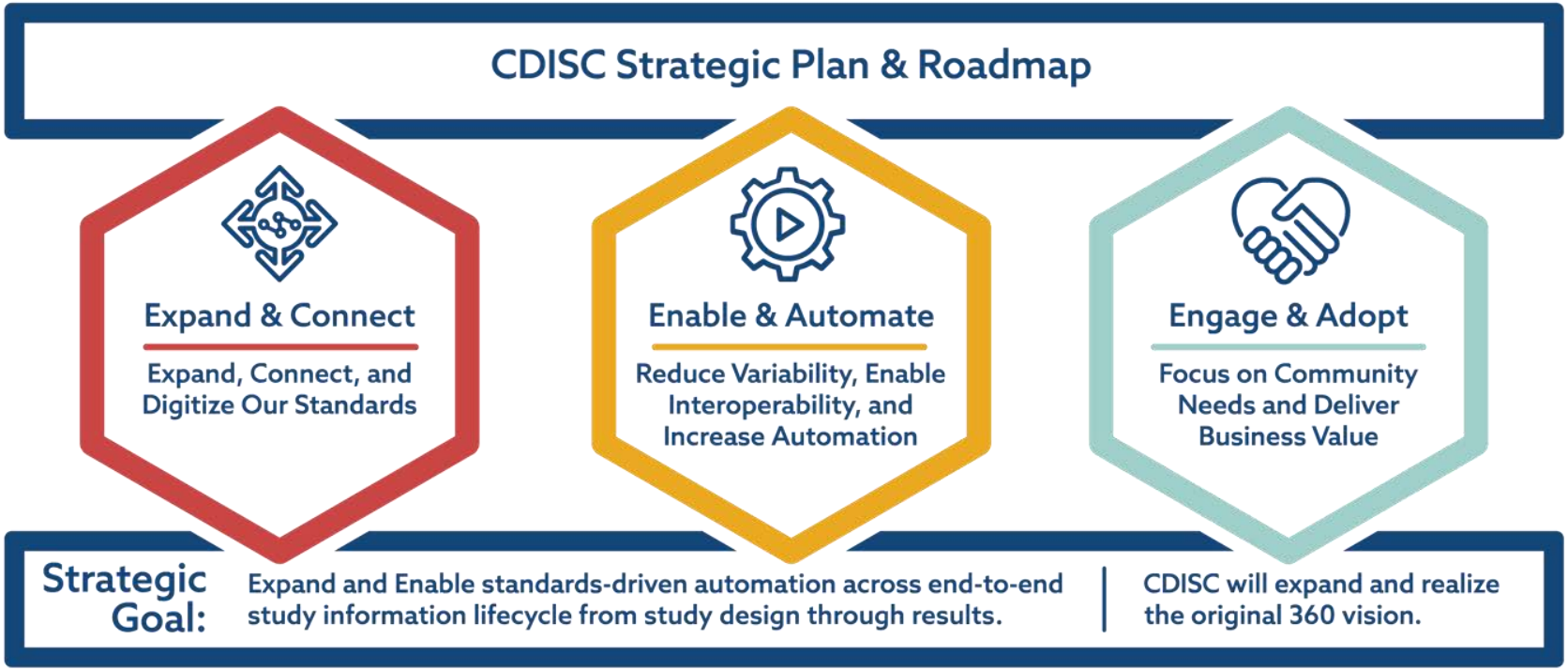


The Result:

- Inefficiency
- Inconsistency
- Lost time and increased cost

Vision for 360i

Realizing the CDISC Mission



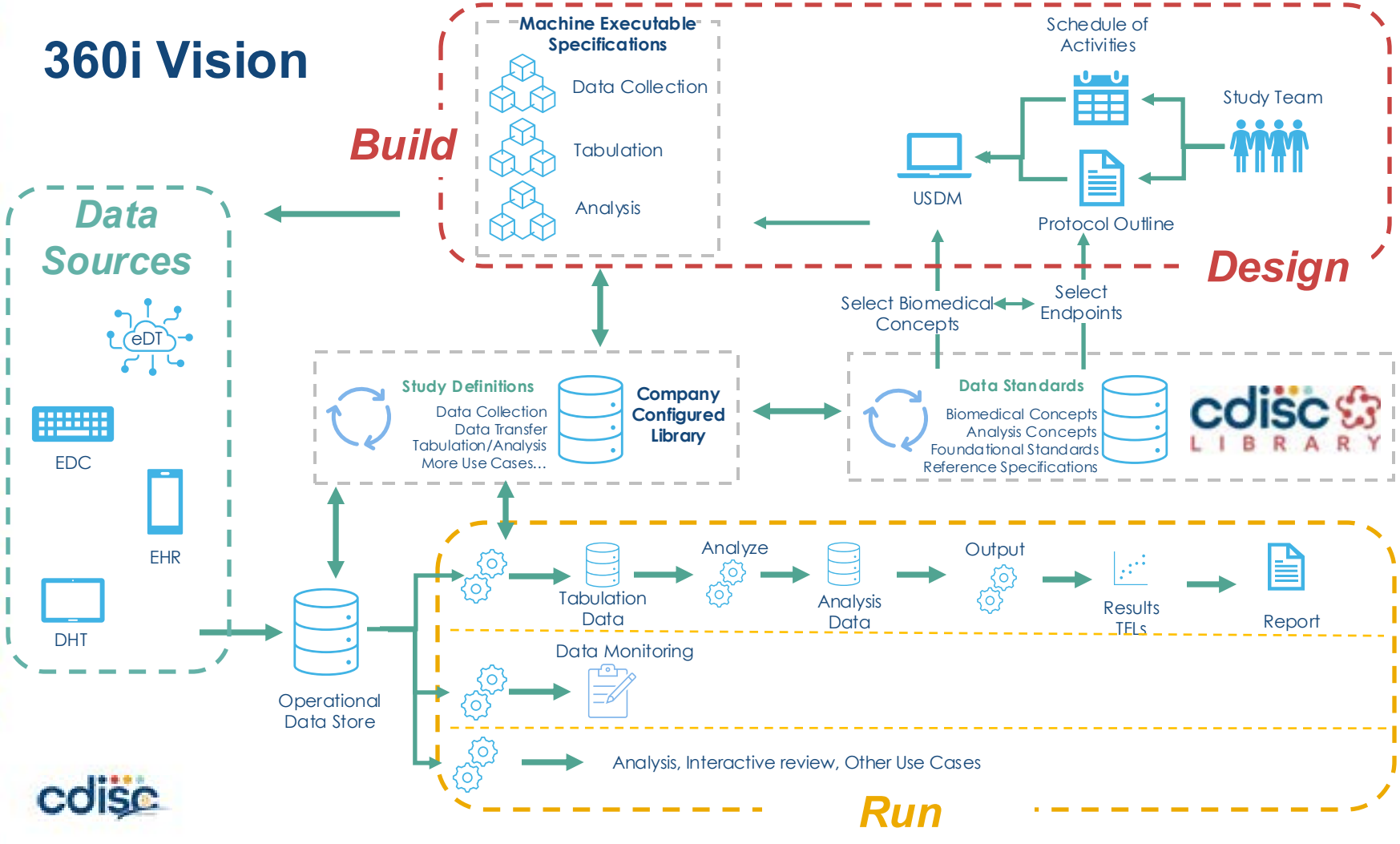
360i Vision

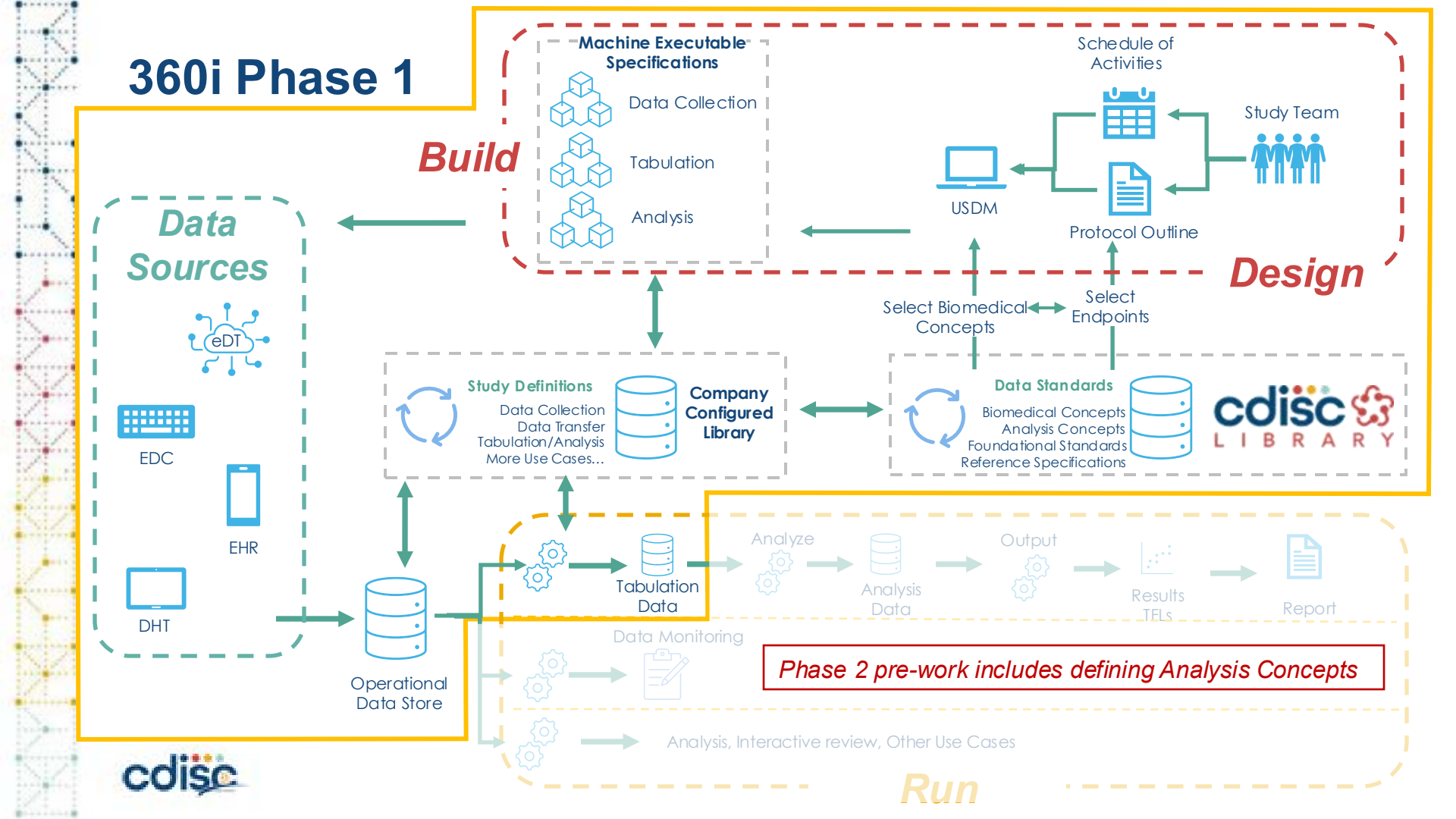
Creating a Digital Ecosystem:

- Unified Study Definitions Model (USDM) as the foundation
- Digital Protocol feed downstream artifacts
- Linked Biomedical Concepts enrich meaning
- Machine-readable, interoperable data flows
- Open-source tools to support transparency and adoption

- **Goal:** Fully digital, standards-driven clinical development lifecycle

360i Vision


Machine Executable Specifications
Data Collection
Tabulation
Analysis
Build
Schedule of Activities
Study Team
USDM
Protocol Outline
Design
Data Sources
eDT
EDC
EHR
DHT
Select Biomedical Concepts
Select Endpoints
Study Definitions
Data Collection
Data Transfer
Tabulation/Analysis
More Use Cases...
Company Configured Library
Data Standards
Biomedical Concepts
Analysis Concepts
Foundational Standards
Reference Specifications
cdisc LIBRARY
Operational Data Store
Tabulation Data
Analyze
Analysis Data
Output
Results TFLs
Report
Data Monitoring
Analysis, Interactive review, Other Use Cases
Run
cdisc

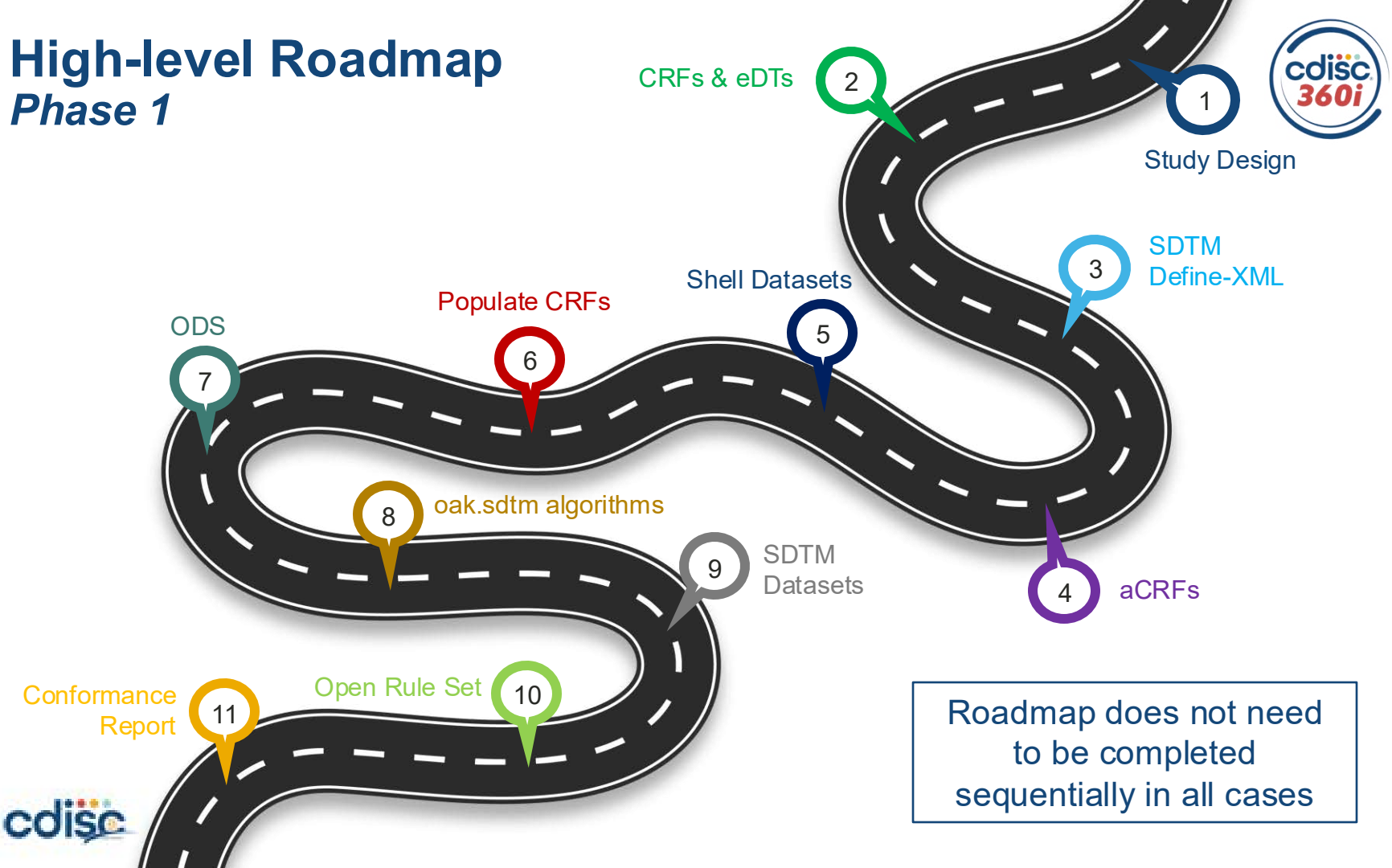

360i Phase 1
Machine Executable Specifications
Data Collection
Tabulation
Analysis
Build
Schedule of Activities
Study Team
USDM
Protocol Outline
Design
Data Sources
eDT
EDC
EHR
DHT
Select Biomedical Concepts
Select Endpoints
Study Definitions
Data Collection
Data Transfer
Tabulation/Analysis
More Use Cases...
Company Configured Library
Data Standards
Biomedical Concepts
Analysis Concepts
Foundational Standards
Reference Specifications
cdisc LIBRARY
Tabulation Data
Analyze
Analysis Data
Output
Results TFLs
Report
Data Monitoring
Operational Data Store
Phase 2 pre-work includes defining Analysis Concepts
Analysis, Interactive review, Other Use Cases
Run
cdisc

AI Innovation

- Engage community to advance standards through innovation
- Open to CDISC Member Organizations
- Solutions may be open-source or commercial
- Open-source encouraged for community impact
- CDISC will not endorse or adopt submitted solutions
- Winners showcased at 2025 US Interchange
- Use Cases
  - Build a USDM-Centric Repository from Legacy Protocols
  - AI-Powered Biomedical Concept Development
  - Predictive Rule Authoring for CORE
  - Reverse Mapping from Analysis to Design via Semantics

360i Technical Roadmap

High-level Roadmap
Phase 1



Roadmap does not need to be completed sequentially in all cases

1. Create Study Design

Design

- LZZT Study Design XLS
- Biomedical Concepts

- USDM JSON Export
- ODM Study Design

Completed (first iteration):

- BC identification & extraction from library
- USDM-JSON generation
- Conformance checks run

To Do:

- Remaining BC identification & extraction
- Remaining conformance checking
- Final USDM-JSON generation

2. Generate CRFs & eDTs

Build

- Study design exports
- Biomedical Concepts
- Dataset Specializations

- ODM-based CRFs
- ODM-based eDTs
- HTML CRFs

Completed (first iteration):

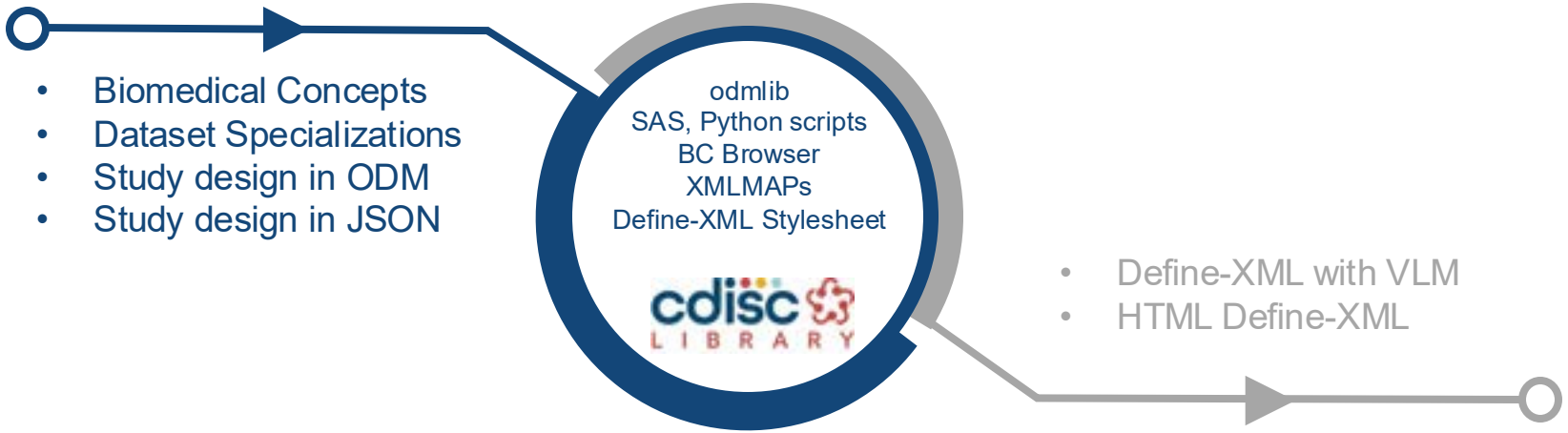
- LZZT study data reverse engineering

To Do:

- ODM-based CRF & eDT generation
- HTML CRF generation

3. Generate SDTM Define-XML

Build



Completed (first iteration):

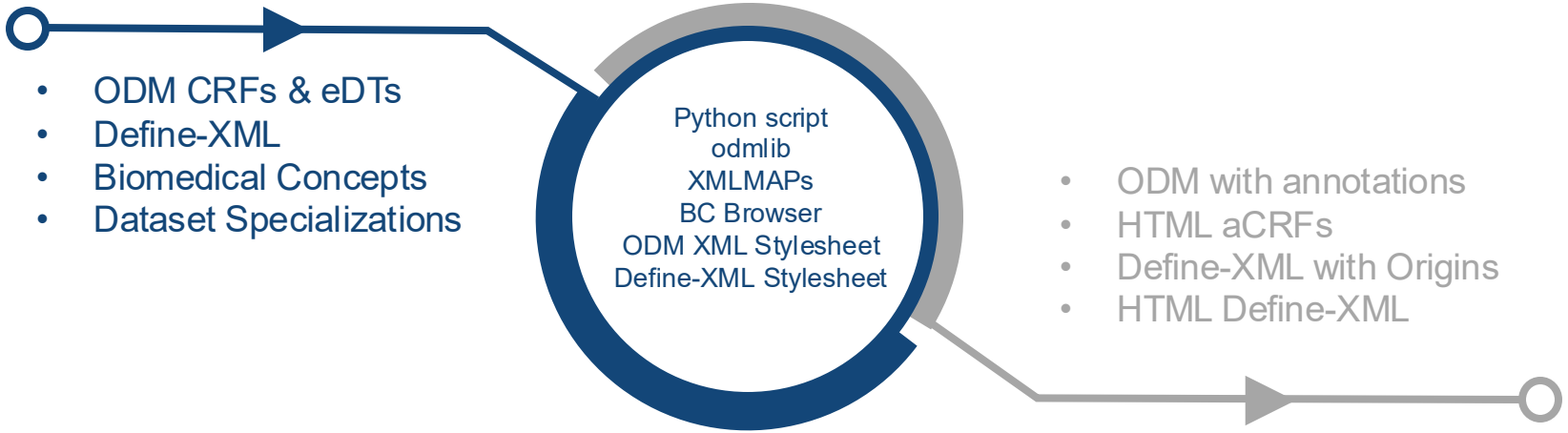
- Partial Define-XML generated
- BCs pulled from USDM JSON export
- CDISC API queried for dataset specializations

To Do:

- Define-XML complete generation with value level metadata

4. Generate aCRFs

Build



Completed (first iteration):

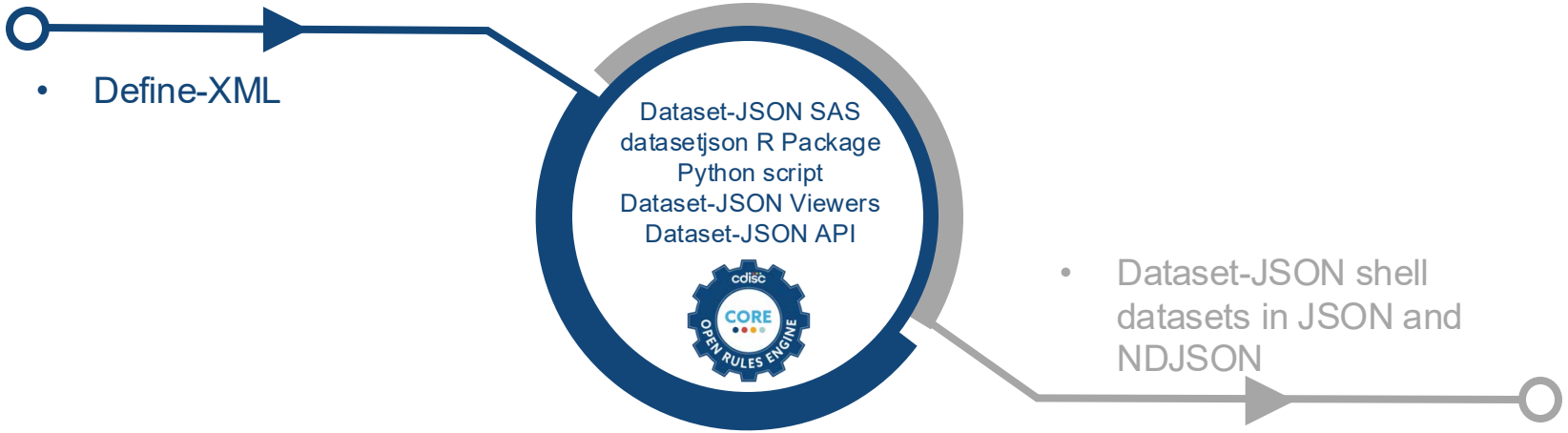
- Partial aCRF generation from BCs
- Partial Define-XML HTML rendering generation

To Do:

- Full ODM & HTML aCRF generation
- Define-XML generation with Origins
- Full Define-XML HTML rendering generation

5. Generate Shell Datasets

Build



Completed (first iteration):

- Partial Dataset-JSON shell generation from biomedical concepts and dataset specializations

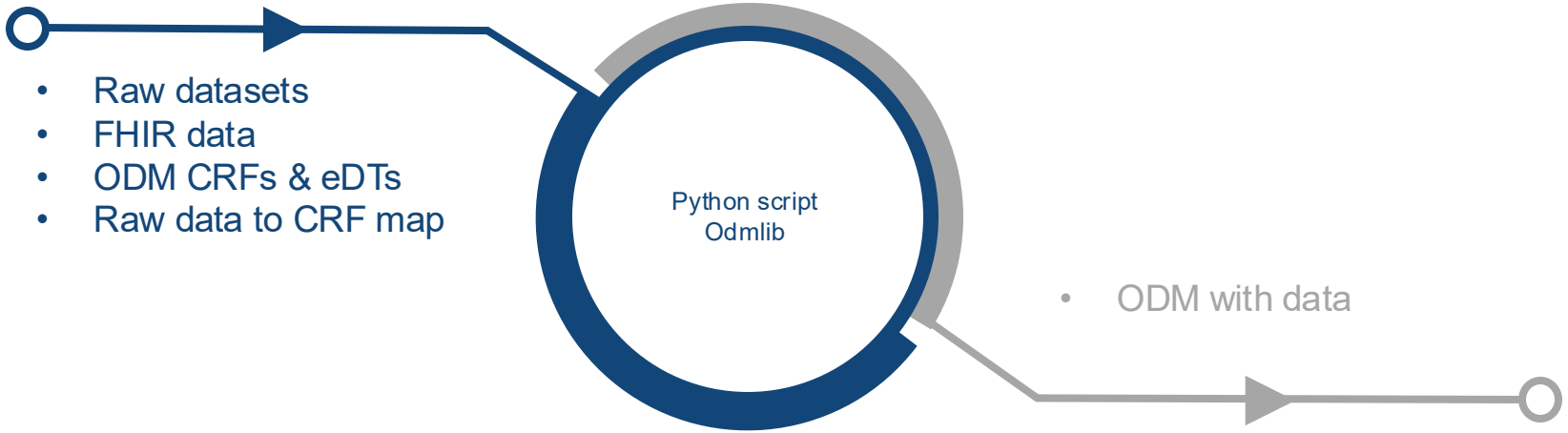
To Do:

- Full Dataset-JSON shell generation

cdisc

6. Populate CRFs & eDTs with Raw Data

Run



Completed (first iteration):

- Reverse engineered raw data obtained
- Synthetic data generated & transformed to SDTM for DM/SU domains

To Do:

- Full ODM generation with CRF and eDT data

7. Load Operational Data Store (ODS)

Run

Completed (first iteration):

- Exploring options for technical infrastructure
- Set up central GitHub repo for project
- Setting up Google Colab notebook to demonstrate the automation

To Do:

- ODS implementation
- Raw data queries

CDISC

8. Create oak.sdtm Algorithm Set

Build

Completed (first iteration):

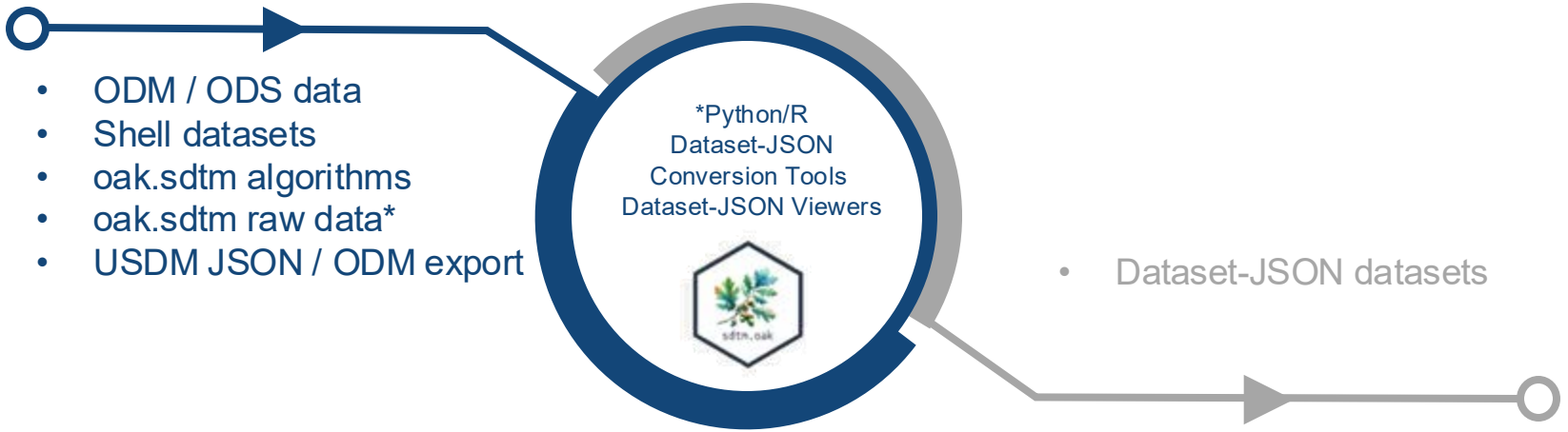
- Partial domain algorithms created

To Do:

- Missing algorithm creation
- Mapping generation

9. Generate SDTM Datasets

Run



Completed (first iteration):

- SDTM Trial design datasets generated

To Do:

- SDTM generation in Dataset-JSON

Bridging the Gap: Automating SDTM Trial Design Datasets Using Structured Protocol

cdisc

cdisc 360i

10. Create Open Rule Set

Build

Completed (first iteration):
- Partial Open Rule generation from BCs

To Do:
- Full Open Rule Set generation

11. Generate Conformance Report

Run

Completed (first iteration):

- None

To Do:

- Full conformance report generation

Technical Roadmap Rules of the Road

Implementation focus is a new way of working that transcends 360i.

Implementation Focus

We'll implement 360i in small iterative increments and learn as we go.

Empirical, Iterative Process

Make it work, make it right, make it fast.

Make it Work

There's more Than One Way To Do It – alternative software tools are welcome.

TMTOWTDI

Make it testable. Continuous Integration / Continuous Deployment.

CI/CD

Pull Requests > Suggestions. What gets worked on gets done.

Just Do It

Benefit from network effects and transparency. Collaborate on GitHub.

Build Publicly

360i Definition of Done

360i has published a complete **pre-configured study package** with all the components defined in **metadata** from study design to submission, **test data** for the study, and **open-source software** to run the study data pipeline to generate analysis results

Thank You!

Join Us in Building the Future:

- Engage with CDISC 360i working groups
- Test, pilot, and provide feedback on new models and tools
- Support the adoption of open digital standards

Together, we are breaking down silos to deliver faster, smarter, more connected clinical research.