cdisc®

2026 Europe Interchange

THE FUTURE IS CONNECTED: STANDARDS AND AI POWERING DIGITAL TRANSFORMATION

MILAN, ITALY | MAIN CONFERENCE: 20-21 MAY | TRAININGS & WORKSHOPS: 18, 19, & 22 MAY

Using Open-Source Tools to Drive Standards-Based Automation From SDTM Setup to CDISC-Compliant Submission

Katarzyna Bzymek, Clinical Data Programmer, SGS Pharma – Clinical Research

Speaker

Katarzyna Bzymek is a Clinical Data Programmer with over three years of experience in clinical research. In her current role, she focuses on implementing and optimizing data standards, developing open-source–driven programming solutions, and exploring innovations in regulatory systems. She is an active contributor to the CDISC CORE project within SGS Pharma.

Katarzyna Bzymek
Clinical Data Programmer
SGS Pharma – Clinical Research

Disclaimer and Disclosures

- *The views and opinions expressed in this presentation are those of the author and do not necessarily reflect the official policy or position of CDISC or SGS Pharma.*

- *The author has no real or apparent conflicts of interest to report.*

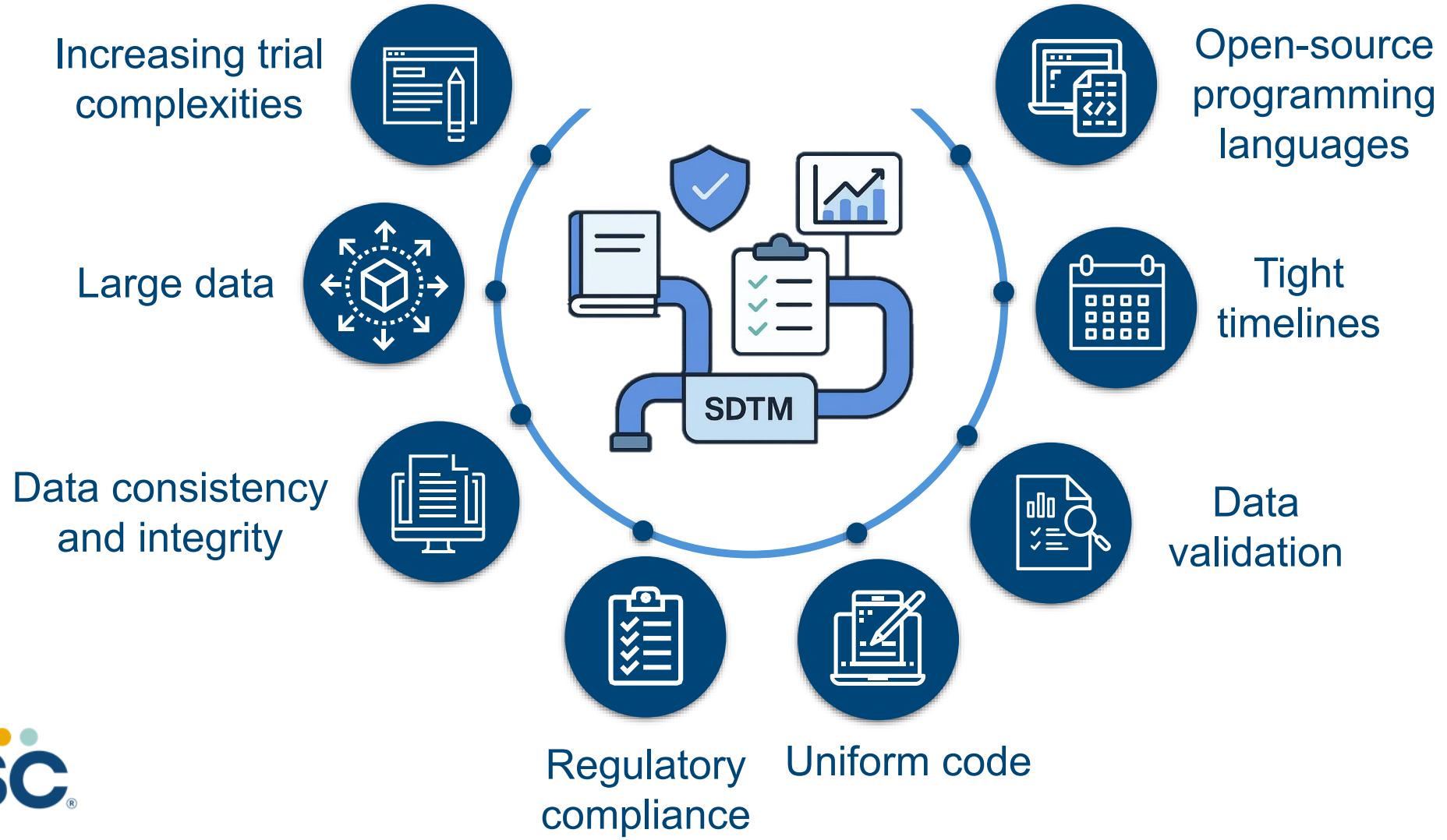
Agenda

1. Introduction
2. Embracing Python in SDTM conversion
3. Study setup workflow
4. Future directions
5. Key takeaways

Introduction

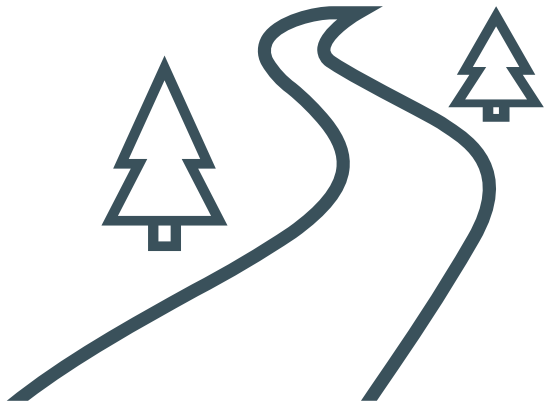
Meeting CDISC requirements in an increasingly complex data landscape

Why SDTM workflows need Smarter tools & Automation Strategies



What we aimed to achieve

- Shorten conversion run times
- Reduce SDTM setup effort
- Efficient code review & version control
- Standardize libraries for aCRF mapping
- Standardize code for derivations

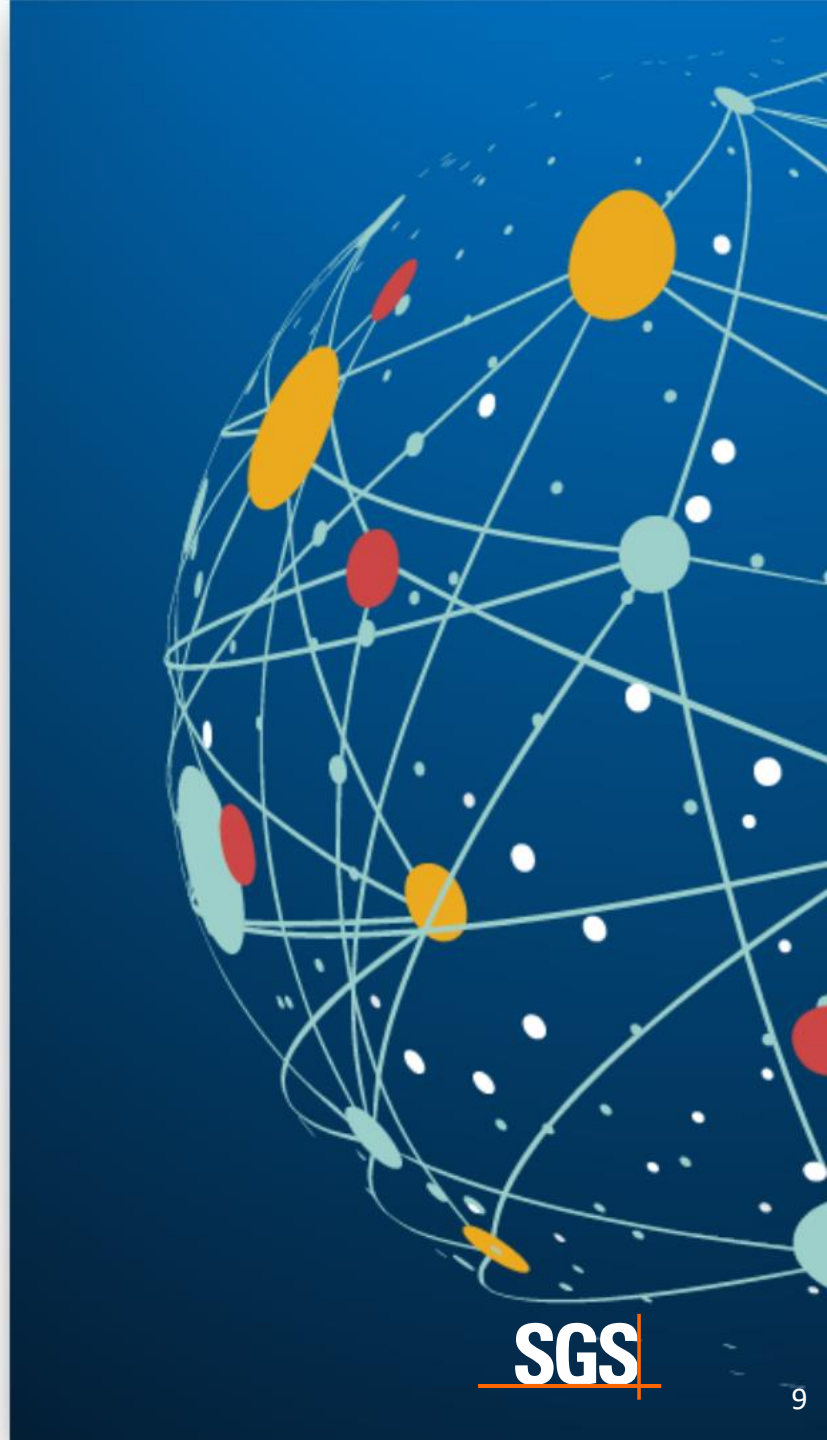


Embracing Python in SDTM Conversion

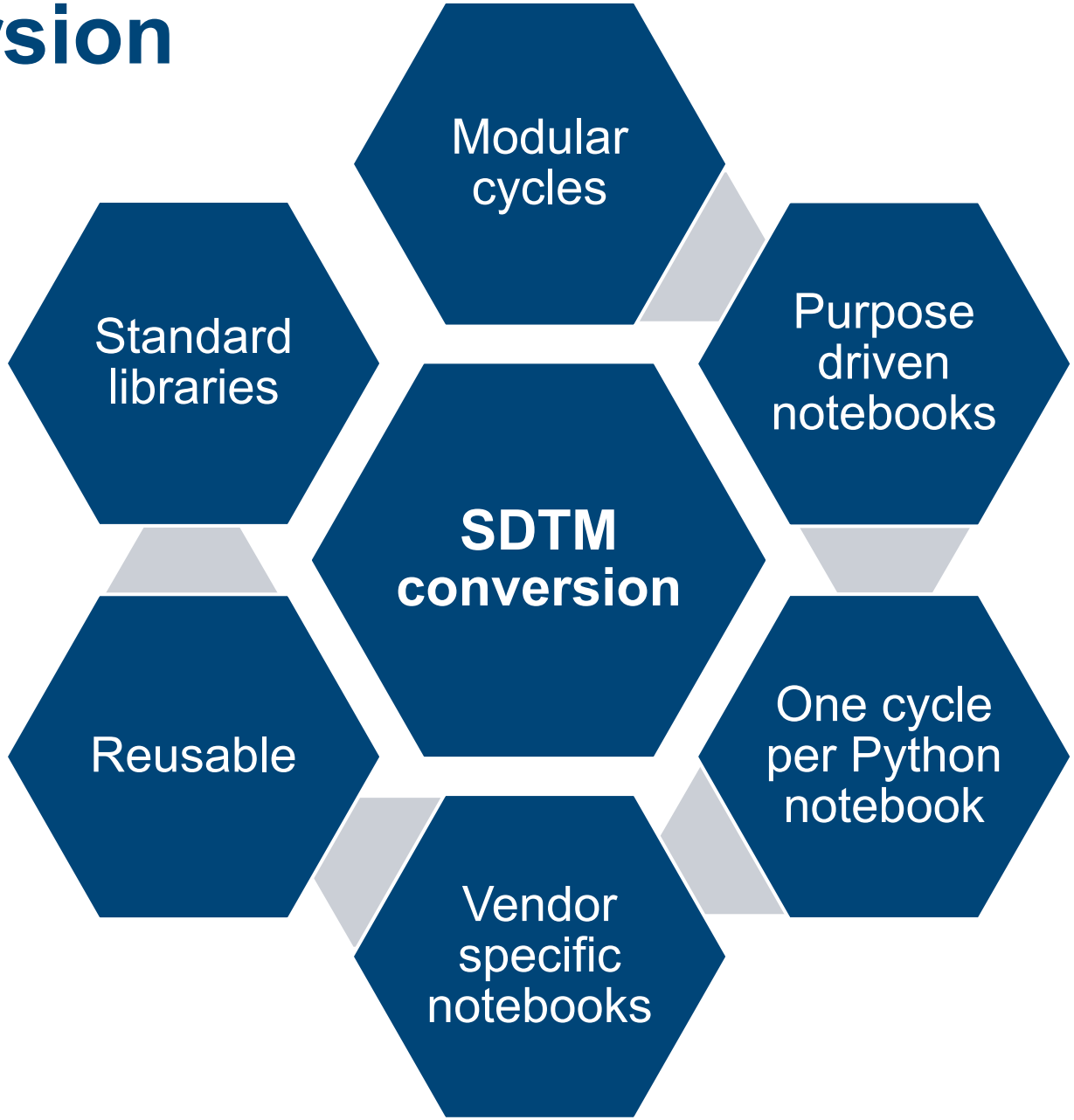
A shift in tools and mindset

Open-Source Foundation: Python

- Core language for SDTM conversion
- Enables reusable components across studies
- Supports automation and standardization
- Supports efficient processing of large datasets (Polars, NumPy)
- Integrated Git and Azure DevOps workflow
- Automated code scan for vulnerabilities in DevOps

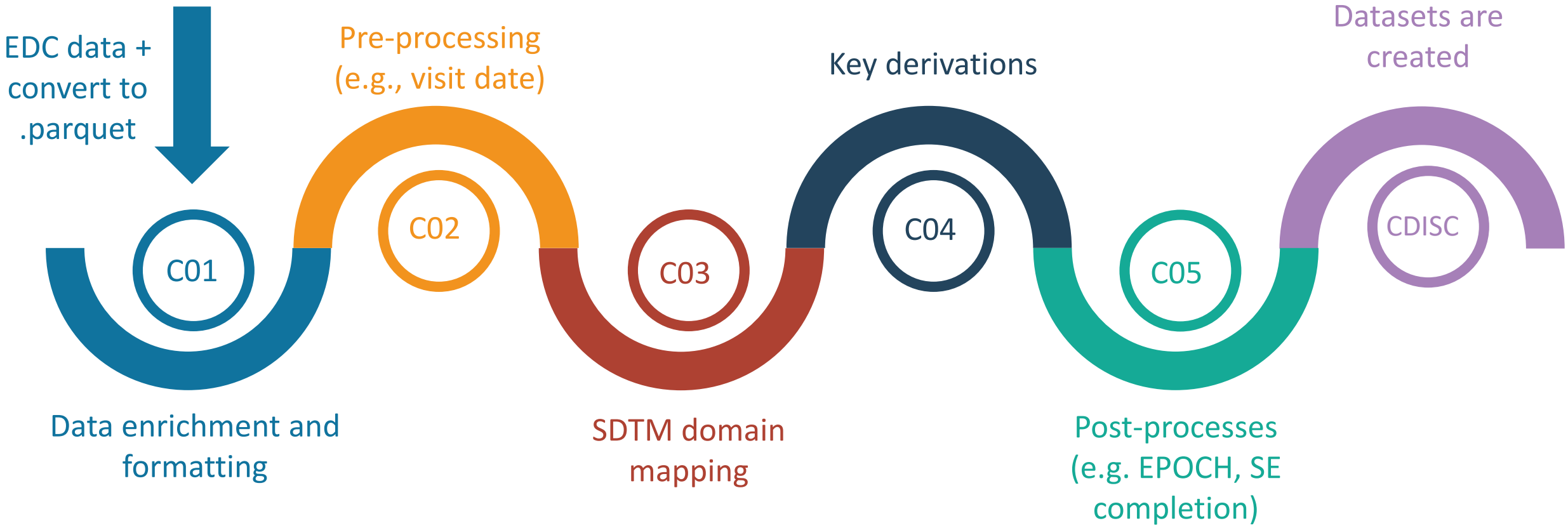


SDTM conversion



SDTM Conversion Workflow

The conversion is structured into reusable, modular cycles.



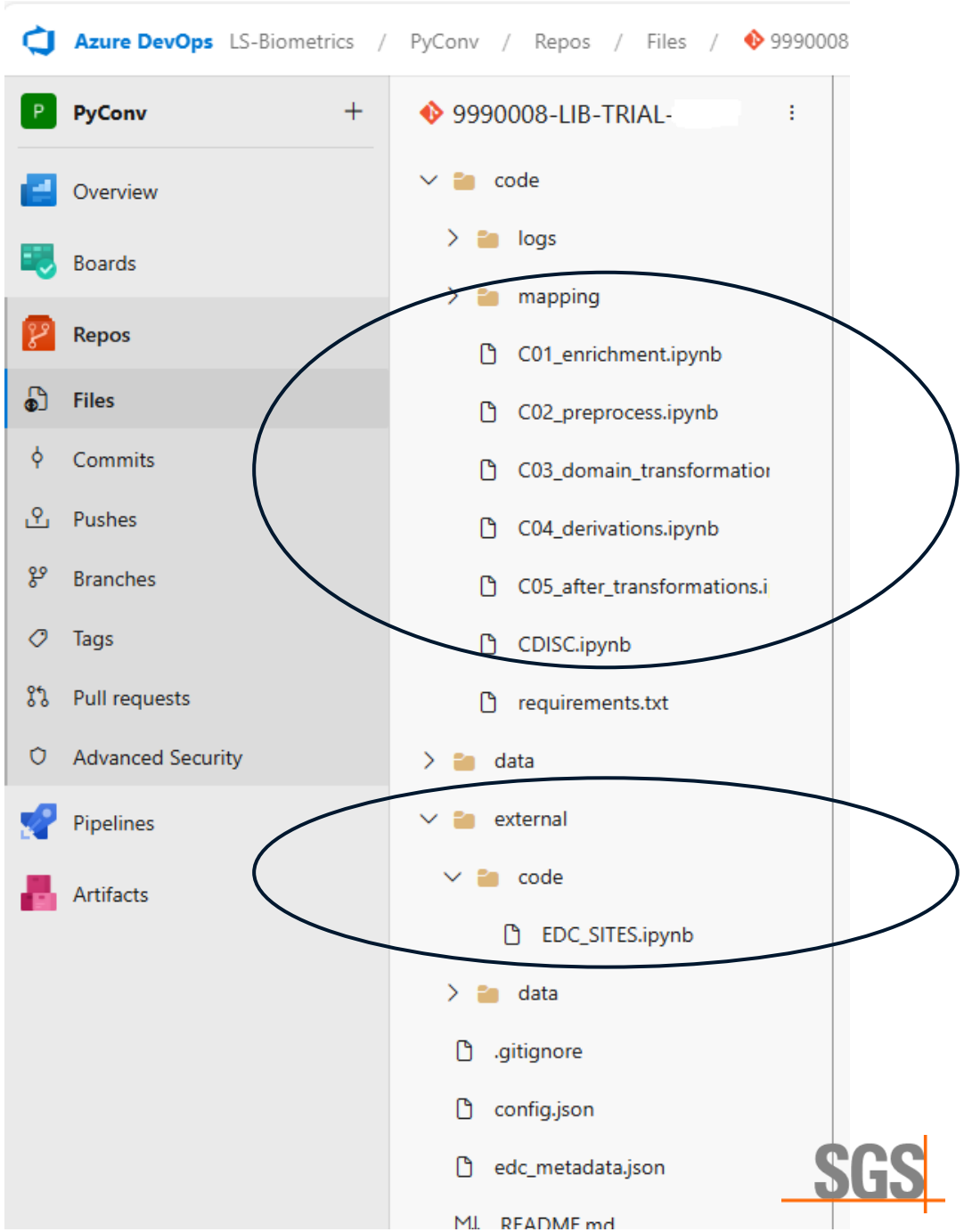
Modular Architecture in Azure Repositories

Conversion notebooks

- Modular cycle notebooks (C01–C05, CDISC)
- Stored and version-controlled in Azure

External vendor data

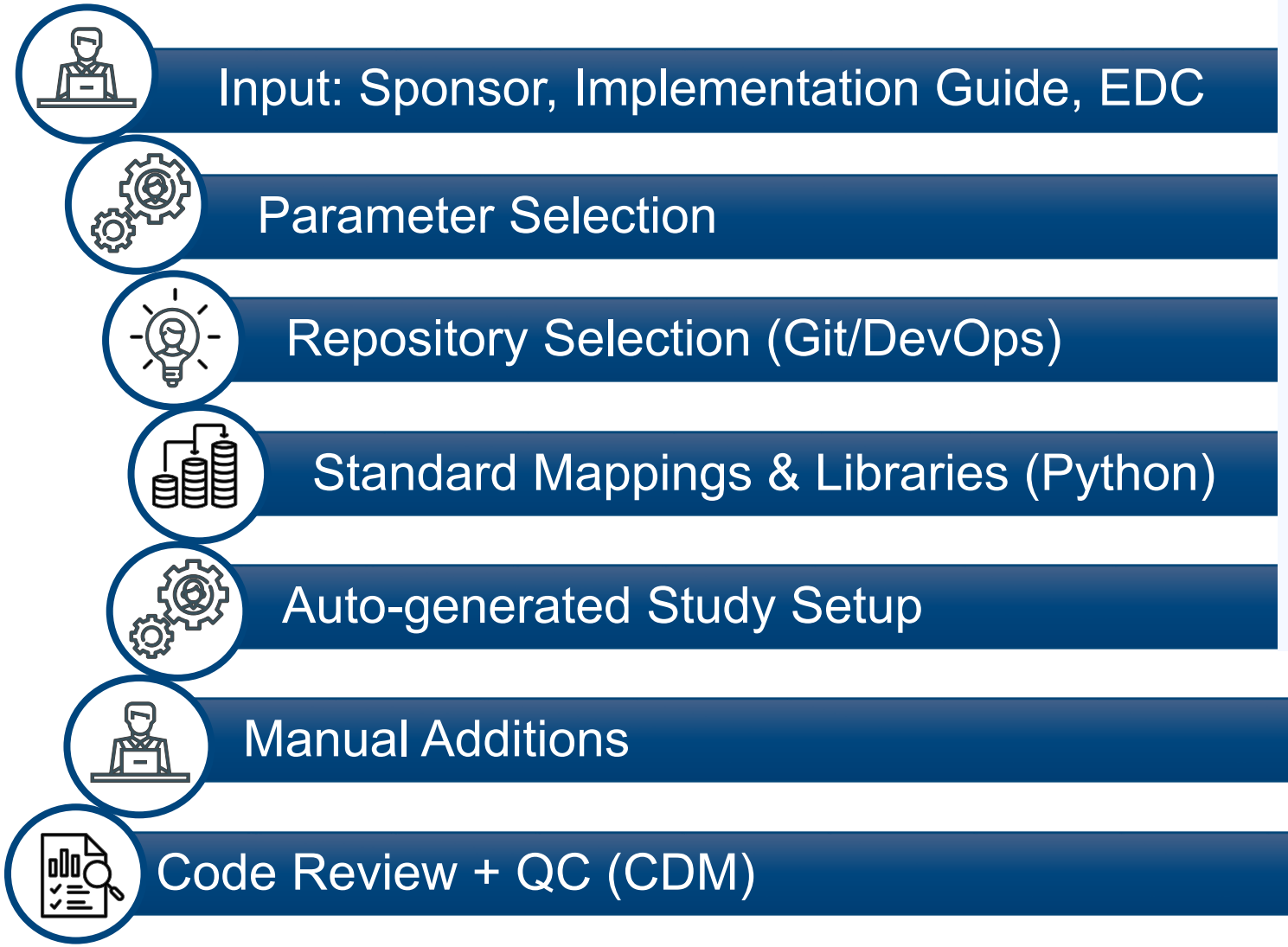
- Separate notebooks
- Upload data via script
- No conversion required

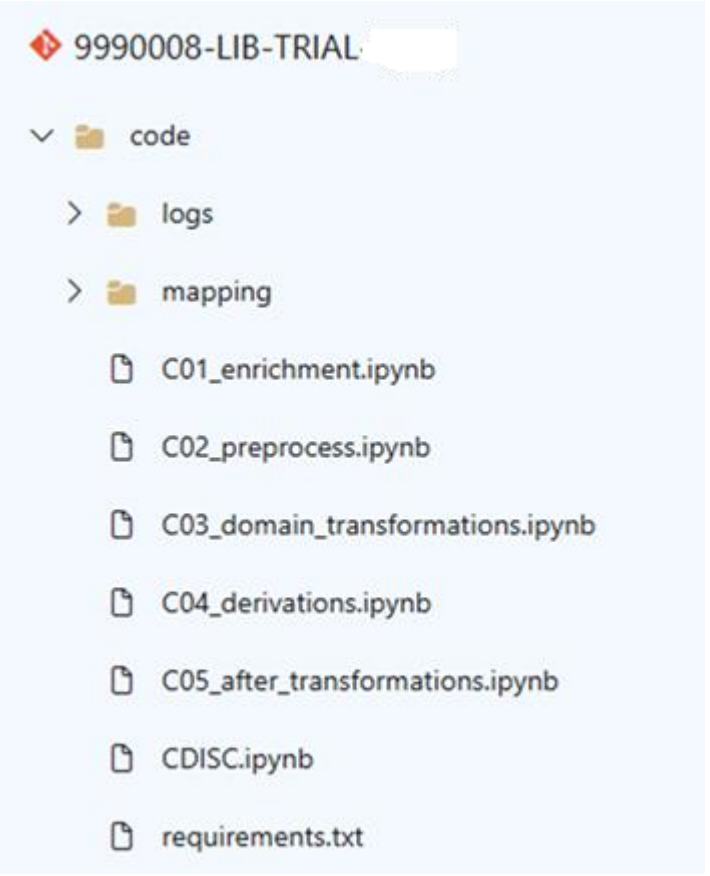


Study setup workflow

Standardized libraries

Study setup workflow



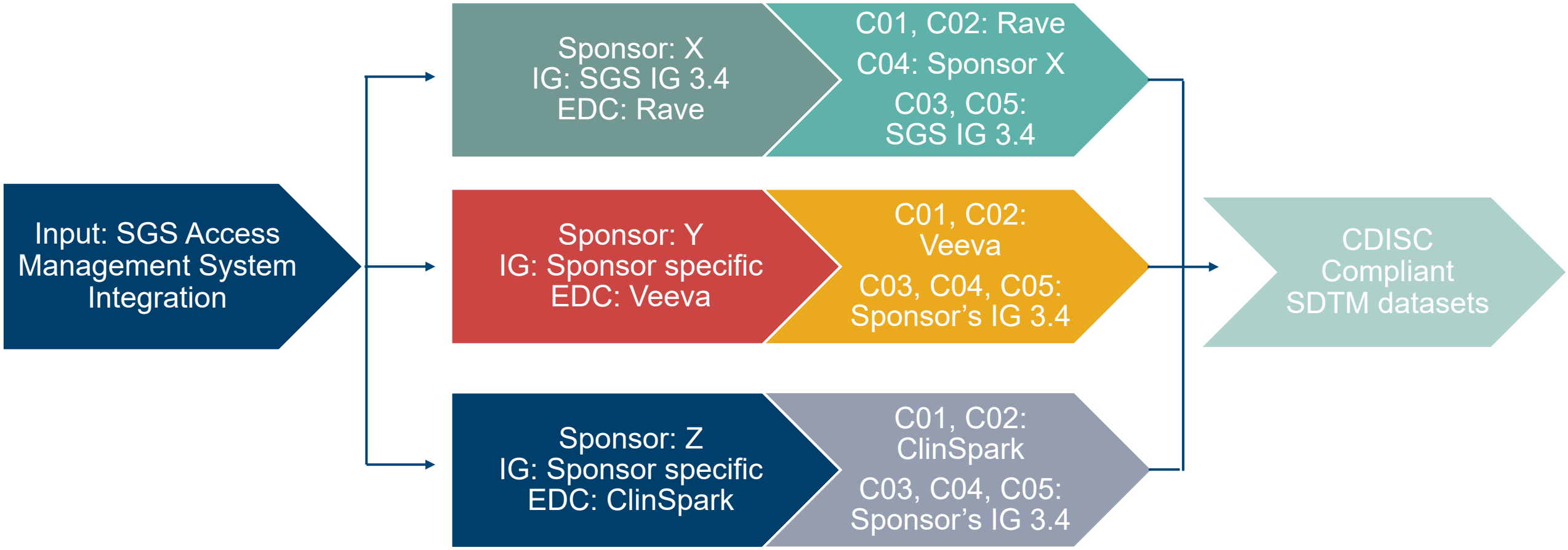


SDTM Libraries: Built for Mix & Match

eDC specific	Implementation guide	Sponsor specific	External data
➢ C01, C02 ➢ Configuration of raw data structure and naming ➢ Reusable across different systems	➢ C03, C04, C05 ➢ Full standard, IG-version specific ➢ Code reviewed and re-used across studies	➢ C04 & C05 ➢ Extends or overrides IG logic only where required ➢ Implements sponsor conventions	➢ Standardized templates per vendor ➢ Aligned to internal SDTM structure

Standard where possible. Custom where required. Controlled by design.

Example: Mix & Match in Practice



Ensuring Standards Integrity Through Version Control

Versioned Libraries

- ✓ All standard & sponsor libraries are centrally maintained and versioned
- ✓ Each study has its repository and traceable library versions

Pre-reviewed standards

- ✓ Standard libraries are code reviewed and approved once
- ✓ Reused across studies without repeated review

Targeted reviews

- ✓ Code reviewers focus on sponsor or study-specific changes/logic
- ✓ Standard Implementation guide logic remains unchanged

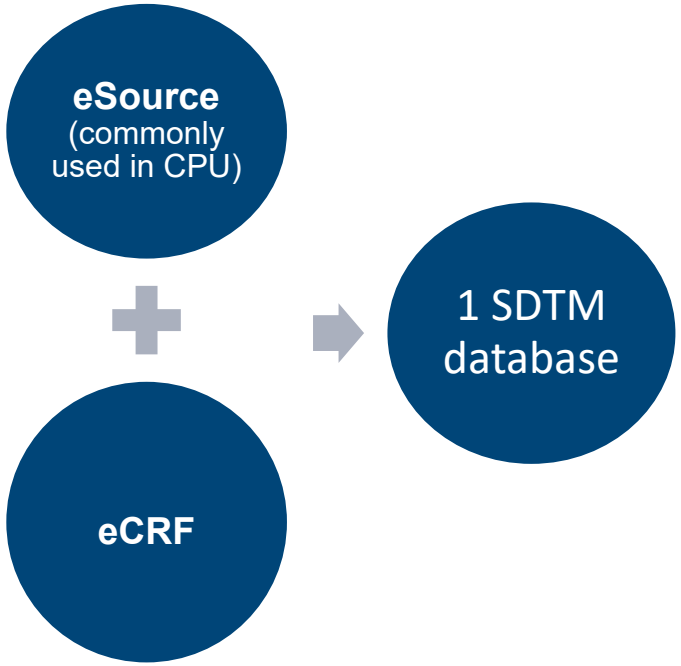
Audit History

- ✓ Full history of what changed, when, why and by who
- ✓ Supports reproducibility and inspection readiness

Future directions

Where to go next…

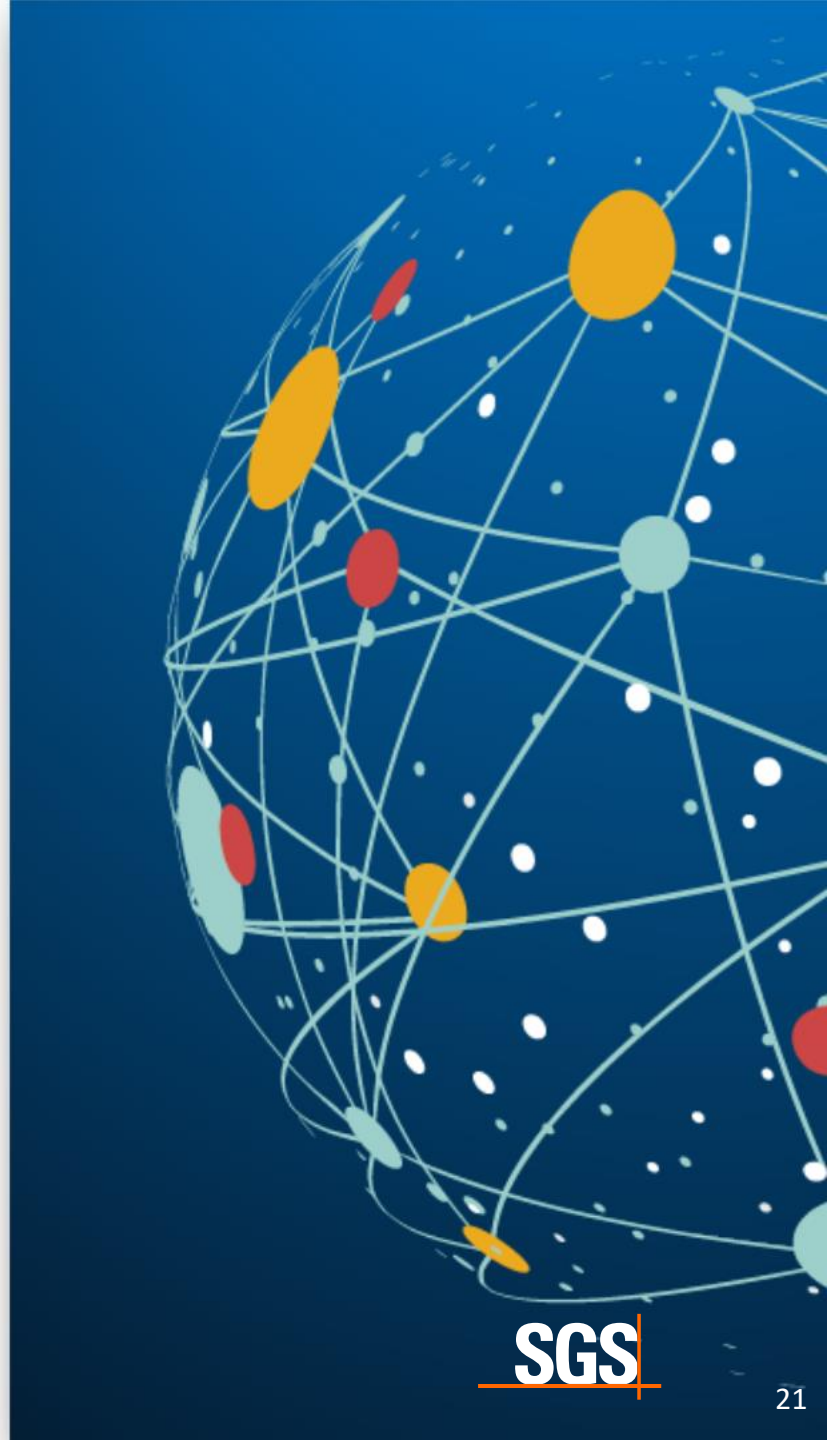
- Further expansion of standardized libraries
- Alignment with evolving CDISC and regulatory expectations
- Metadata automation
- CDISC 360i readiness
- Different EDC systems ⟶ one standardized SDTM conversion



Key takeaways

Key takeaways

- Modular conversion design enables reusable SDTM workflows
- Parameter-driven setup reduces manual effort
- Standard libraries improve consistency and quality
- Code review is performed once and reused across studies
- Open-source tools enable flexibility and scalability



Thank You!

Contact us

clinicalresearch@sgs.com

www.sgs.com/pharma

sgs.com/linkedinpharma

Visit us at the SGS Booth