

PhUSE EU CONNECT 2018

SI10

Implementation of a Metadata-based Approach to Statistical Planning, Analysis & Reporting

By Frank Freischläger & Hanspeter Schnitzer



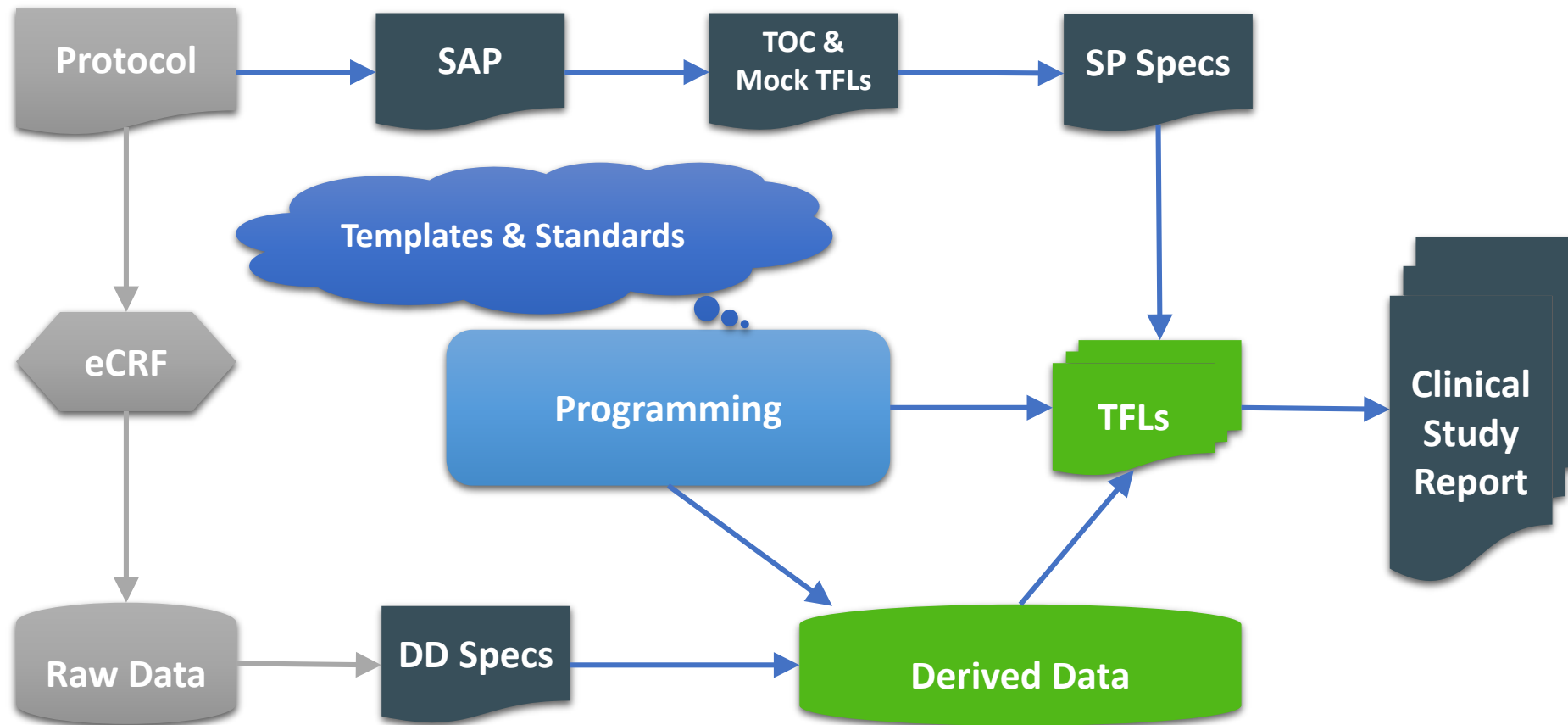
INDEX

Situation & Purpose
Clinical Study Metadata
Statistical Programming for Analysis
Statistical Reporting
Statistical Planning
Conclusion



SITUATION without METADATA

Traditional approach of statistical planning, analysis & reporting of clinical studies



PURPOSE



- Metadata approach to biostatistics
- Consider all aspects of processing
- Set up a system built on and compliant with standards
- Assure flexibility in support of clients and sponsors
- Let the system do the routine work
- Limit needs for programming to a minimum
- Allow statisticians and programmers to concentrate on advanced methods
- Enable respective shift in manpower

CLINICAL STUDY METADATA



METADATA REPOSITORY

Overview of data panels with content explicitly required per project

General

- Clients
- Stakeholders
- Projects
- Studies

Protocol

- Documents
- IE Criteria
- Objectives
- Analyses

Design

- Epochs
- Visits
- Disposition
- Exposure

Data

- Sources
- Groups
- Variables

Deliverables

- Plans
- Derived datasets
- Outputs
- Reports

Statistics

- Populations
- Endpoints
- Intercurrent events
- Estimands
- Evaluations
- Methods
- References

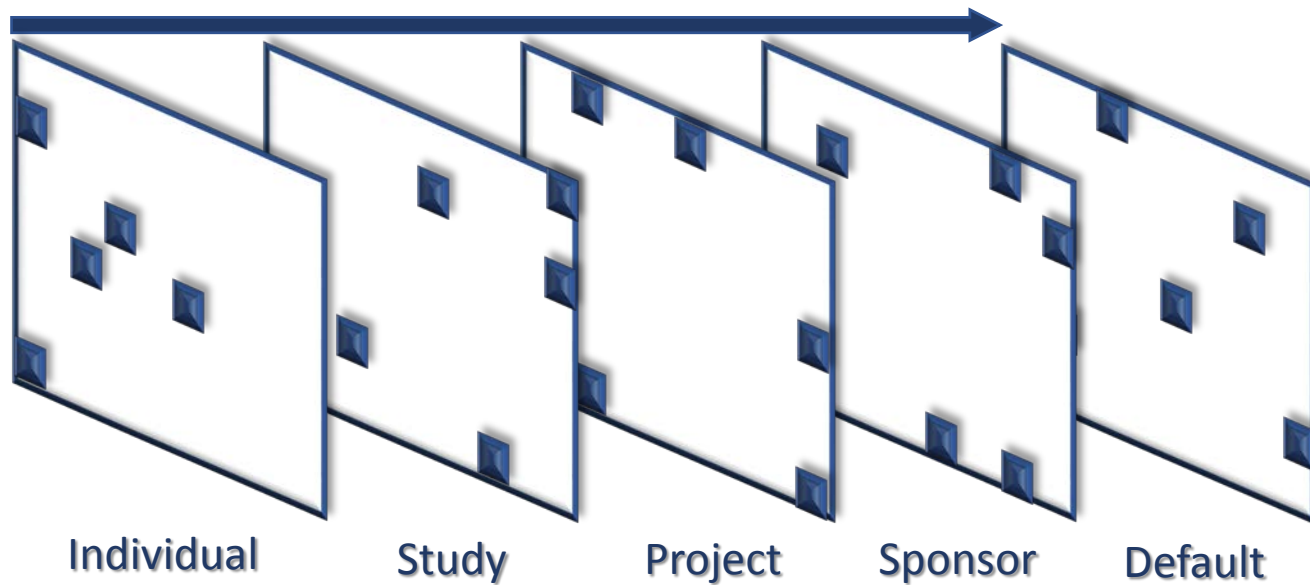
Programming

- LibDirSpecs
- Output structures
- Report structures
- Plan structures

FLEXIBILITY in APPLYING STANDARDS

Approach with different layers

- Project-specific content is mandatory
- Other content may be defined on higher levels
- Levels are dynamic
- Content is guaranteed
- Metadata entry or import
- Metadata edit checks are in place
- Independent double entry as an option



METADATA REPOSITORY

Overview of data panels with content that may be given on any level

Statistics

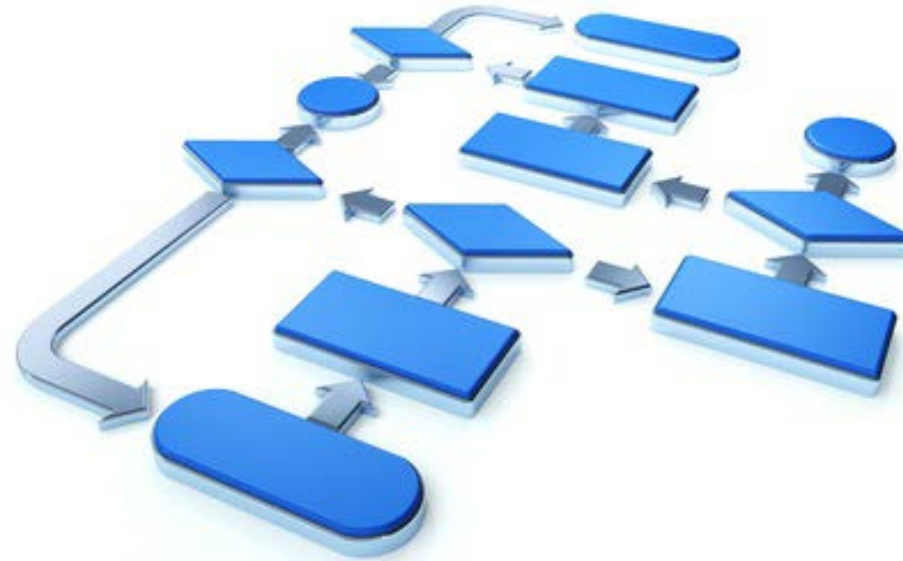
- Semantics
- Tools
- Templates
- Method options
- Derivations
- Descriptors
- Time definitions
- Date imputations

Programming

- InDataset Specs
- Settings
- Styles
- Output options
- Report options
- Plan options

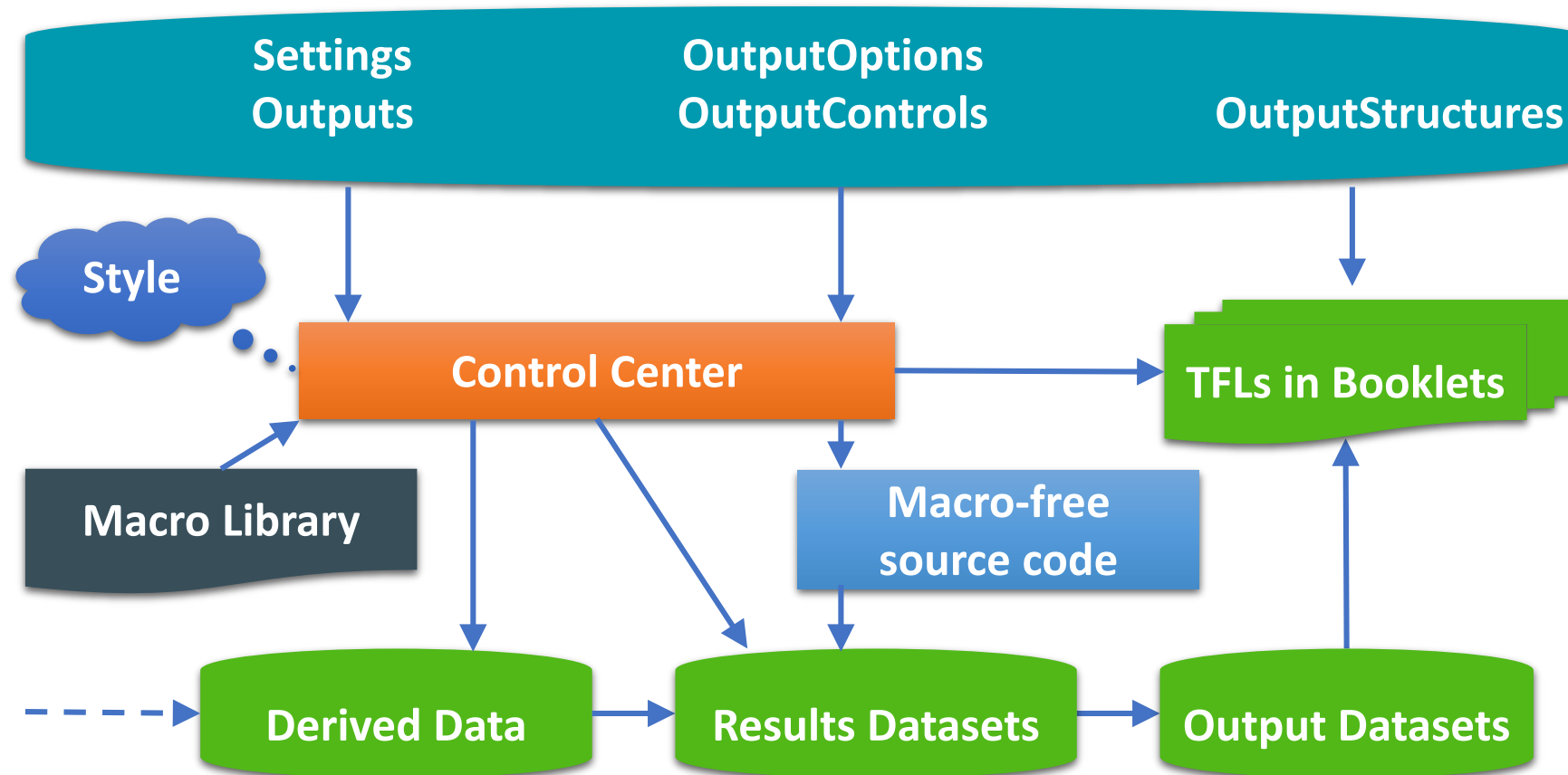
- 🌐 Resulting in project relevant metadata
- 🌐 Processed and archived per production and delivery

STATISTICAL PROGRAMMING for ANALYSIS



WORKFLOW

From derived datasets to booklets of outputs



- For development programming
- With original SAS outputs
- With log checking summaries

LAYOUTS and FORMATTING

Example of PhUSE CSS Analyses and Code Sharing WG on demographics

Table 7.1. Demographic Summary
<Insert population (for example, Safety Population (N = xxx))>
<Insert study ID(s) or description of database utilized
Study Phase or phases (if needed)>

Demographic Parameter		PL (N=xxx)	T1 (N=xxx)
Sex n (%)	n*	xx	xx
	Female	xx (x.x)	xx (x.x)
	Male	xx (x.x)	xx (x.x)
	Missing	xx	xx
Age (years)	n*	xx	xx
	Mean	xx.x	xx.x
	SD	xx.x	xx.x
	Median	xx.x	xx.x
	Q1, Q3	xx, xx	xx, xx
	Min, Max	xx, xx	xx, xx
	Missing	xx	xx
Age Categories n (%)	n*	xx	xx
	<65	xx (x.x)	xx (x.x)
	≥65 and <75	xx (x.x)	xx (x.x)
	≥75 and <85	xx (x.x)	xx (x.x)
	≥85	xx (x.x)	xx (x.x)
	Missing	xx	xx
	≥65	xx (x.x)	xx (x.x)
	≥75	xx (x.x)	xx (x.x)

Label, n (%)
Label - n (%)
Label n (%)

Label, Unit
Label (Unit)
Label [Unit]

Missing category required?

xx (xx.x)
xx (xx.x)
xx (xxx.x)

If 100%?
N (100.0)
N (100)

If 0?
0 (0.0)
0
-
.

STATISTICAL REPORTING



DATA-TO-TEXT TECHNOLOGY

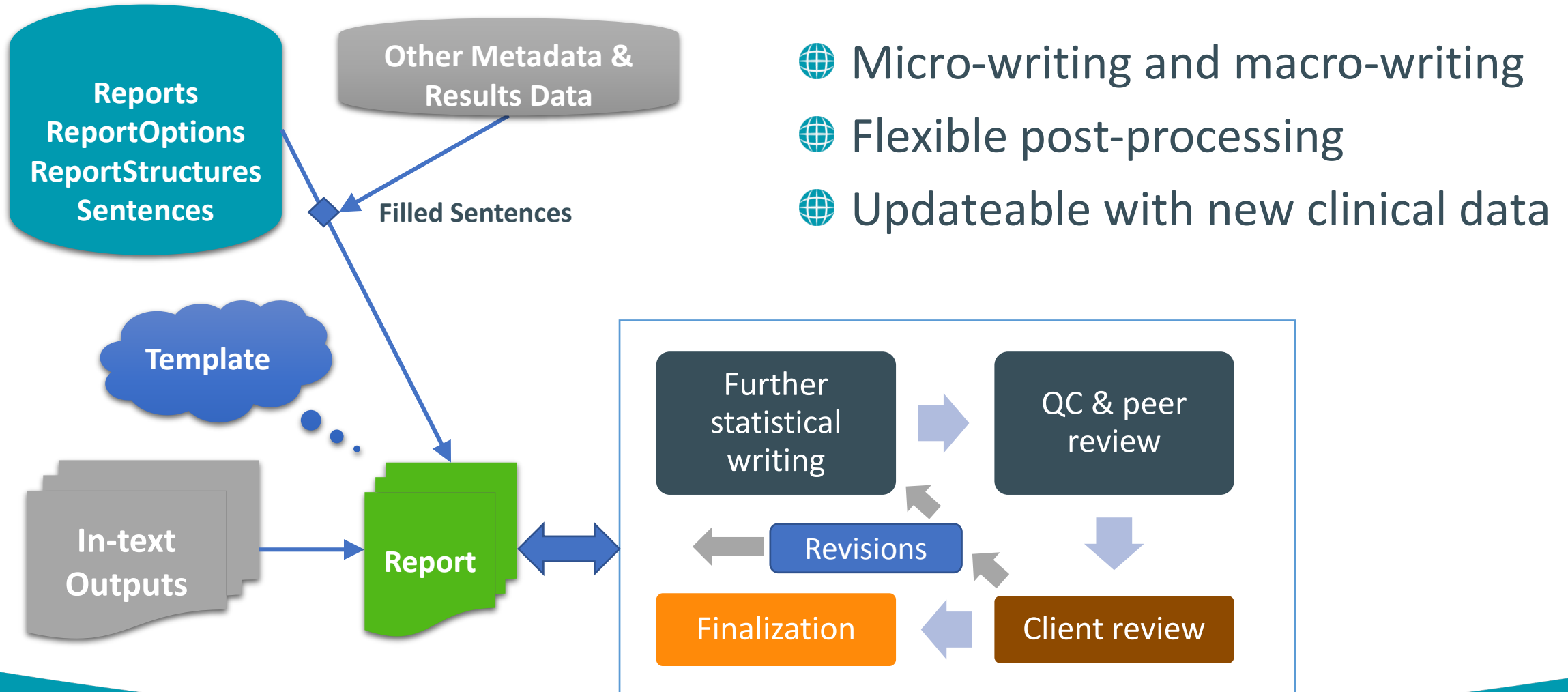
Utilizing methods of Natural Language Generation

- 🌐 Developed by Ehud Reiter and others
- 🌐 From structured data to text messages and documents
- 🌐 Automatic creation of sentences and documents
- 🌐 Content selection and surface realization in computational linguistics
- 🌐 Examples include weather forecasts, automated journalism
- 🌐 Published successful applications to healthcare data



WORKFLOW

Auto-generated report elements and post-processing



SENTENCES for a REPORT

Some examples: opening line & short description of demographics

A total of 500 patients were screened, of whom 50 (10.0%) failed screening prior to randomization (see Table 1).

A total of [DISP_SCR_TOT_N] [SUBJ_TXT] were screened, of whom [DISP_FAIL_TOT_N] ([DISP_FAIL_TOT_PCT] %) failed screening prior to randomization ([REF_TXT1] Table [REF_PT_DISP]).

Most patients were female (62.2%) and white (54.4%).

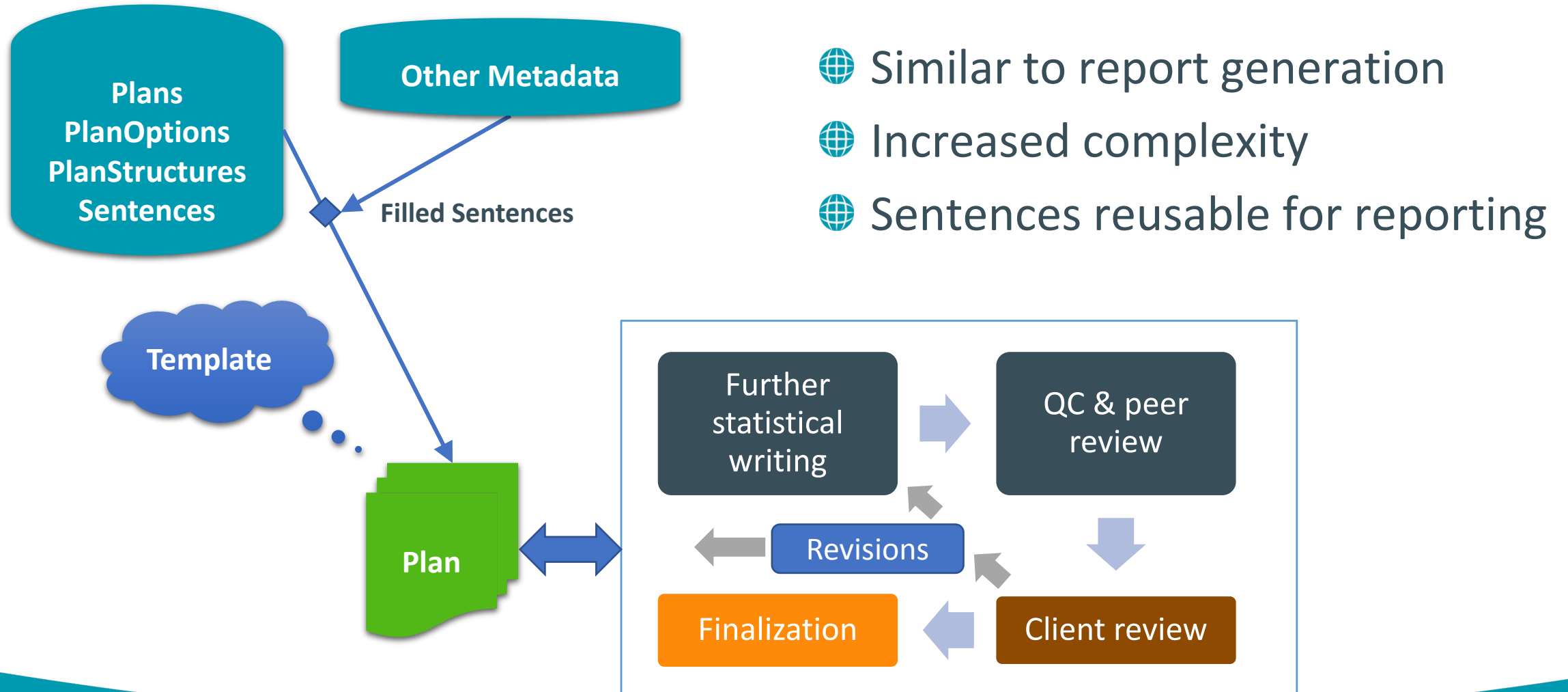
Most [SUBJ_TXT] were [DEMO_GENDER_MOSTFRQ_LBL] ([DEMO_GENDER_MOSTFRQ_PCT] %) and [DEMO_RACE_MOSTFRQ_LBL] ([DEMO_RACE_MOSTFRQ_PCT] %).

STATISTICAL PLANNING



WORKFLOW

Auto-generated elements of a planning document, and post-processing



SENTENCES for a PLAN

Some examples: definition of TEAE & description of an estimand

An adverse event is treatment-emergent, if onset or deterioration of the event appear after the first study treatment intake and no later than after the last study treatment.

An adverse event is treatment-emergent, if onset or deterioration of the event appear after the first [TRT_LBL] [TRTAPPL_LBL] and no later than [TE_LAGDEF] [TE_LAGUNIT] after the last [TRT_LBL] [TRTAPPL_LBL] .

Population-level summary

Measure

Difference in means between treatment conditions in the change from baseline to Week 24 in sBP in the targeted population regardless of intercurrent events.

Parameter

Patient population

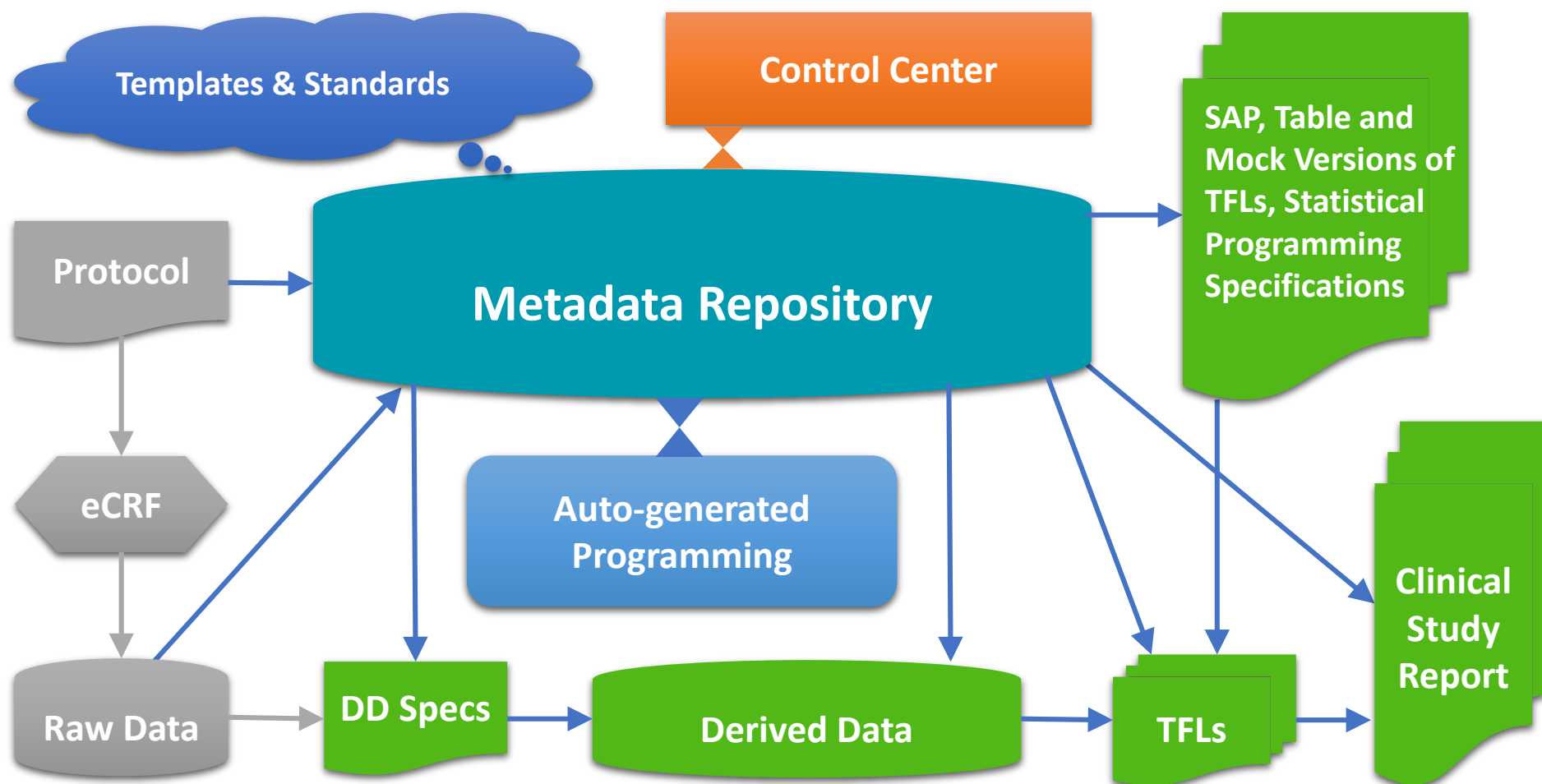
Specification of consideration of intercurrent events

CONCLUSION



METADATA-BASED APPROACH

New approach of statistical planning, analysis & reporting of clinical studies



SUMMARY

- 🌐 Extensive MDR
- 🌐 Flexible & neural
- 🌐 Productive & reliable
- 🌐 Requires careful metadata management
- 🌐 Limitations
- 🌐 Challenging in preparing coded links for sentences

OUTLOOK

- 🌐 Applicable to quality documents like SOPs etc.
- 🌐 Useful for service proposals
- 🌐 Improve MDR frontend, backend and control center
- 🌐 Consider OpenXML or Linked Data as alternatives





THANK YOU



Frank Freischläger

Estimondo GmbH

frank.freischlaeger@estimondo.com

www.estimondo.com

Hanspeter Schnitzer

Estimondo GmbH

hanspeter.schnitzer@estimondo.com

www.estimondo.com