Using SAS® Macros to Simplify Preparation of SDTM Data, Annotated CRFs and Define.xml

PhUse 2009, Basel

Niels Both
Principal Consultant
S-Cubed

Presentation - Overview

eCTD

Overall Architecture

Define.xml

Annotated CRF

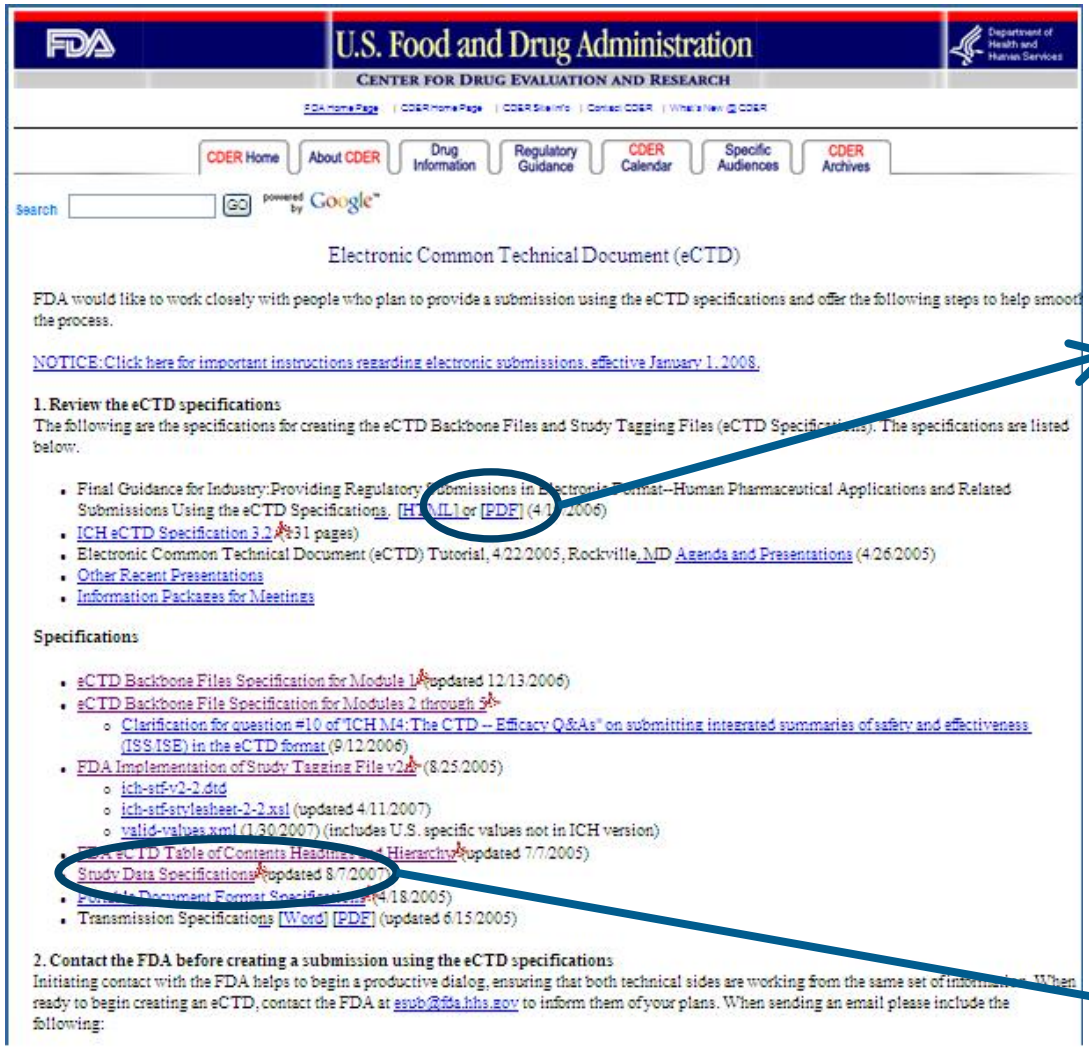
SDTM Generation

SDTM Step 1: Mapping

SDTM Step 2: Generic Macros

Conclusion

Background - eCTD



U.S. Food and Drug Administration

CENTER FOR DRUG EVALUATION AND RESEARCH

CDER Home | About CDER | Drug Information | Regulatory Guidance | CDER Calendar | Specific Audiences | CDER Archives

Search

Electronic Common Technical Document (eCTD)

FDA would like to work closely with people who plan to provide a submission using the eCTD specifications and offer the following steps to help smooth the process.

NOTICE: Click here for important instructions regarding electronic submissions, effective January 1, 2008.

1. Review the eCTD specifications

The following are the specifications for creating the eCTD Backbone Files and Study Tagging Files (eCTD Specifications). The specifications are listed below.

- Final Guidance for Industry: Providing Regulatory Submissions in Electronic Format--Human Pharmaceutical Applications and Related Submissions Using the eCTD Specifications, [HTML] or [PDF] (4/1/2006)
- ICH eCTD Specification 3.2 (31 pages)
- Electronic Common Technical Document (eCTD) Tutorial, 4/22/2005, Rockville, MD Agenda and Presentations (4/26/2005)
- Other Recent Presentations
- Information Packages for Meetings

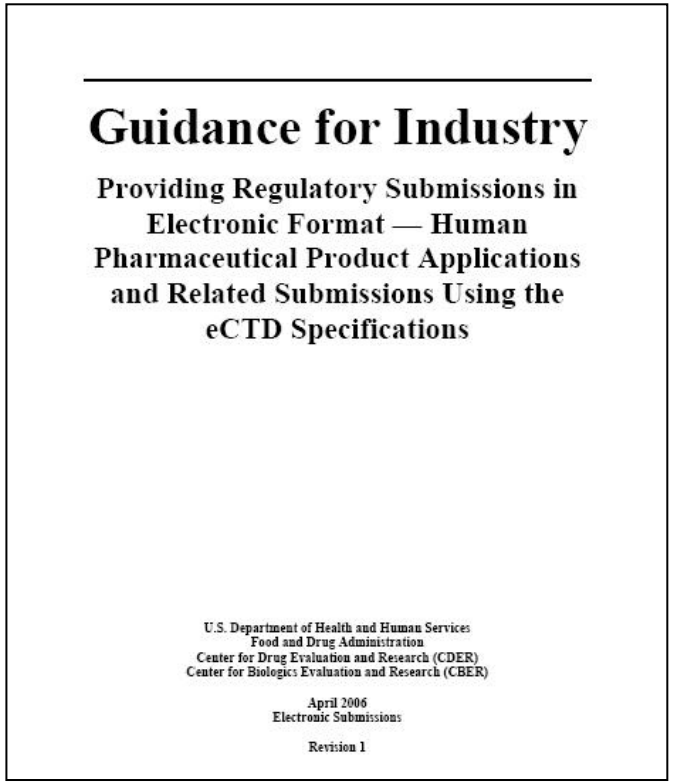
Specifications

- eCTD Backbone Files Specification for Module 1 (updated 12/13/2006)
- eCTD Backbone File Specification for Modules 2 through 5
  - Clarification for question #10 of ICH M4: The CTD -- Efficacy Q&As" on submitting integrated summaries of safety and effectiveness (ISS ISE) in the eCTD format (9/12/2006)
- FDA Implementation of Study Tagging File v2 (8/25/2005)
  - ich-stf-v2-2.dtd
  - ich-stf-stylesheet-2-2.xsl (updated 4/11/2007)
  - valid-values.xml (1/30/2007) (includes U.S. specific values not in ICH version)
- FDA eCTD Table of Contents Headings and Hierarchy (updated 7/7/2005)
- Study Data Specifications (updated 8/7/2007)
- Provide Document Format Specifications (4/18/2005)
- Transmission Specifications [Word] [PDF] (updated 6/15/2005)

2. Contact the FDA before creating a submission using the eCTD specifications

Initiating contact with the FDA helps to begin a productive dialog, ensuring that both technical sides are working from the same set of information. When ready to begin creating an eCTD, contact the FDA at esub@fda.hhs.gov to inform them of your plans. When sending an email please include the following:

http://www.fda.gov/cder/regulatory/ersr/ectd.htm



Guidance for Industry

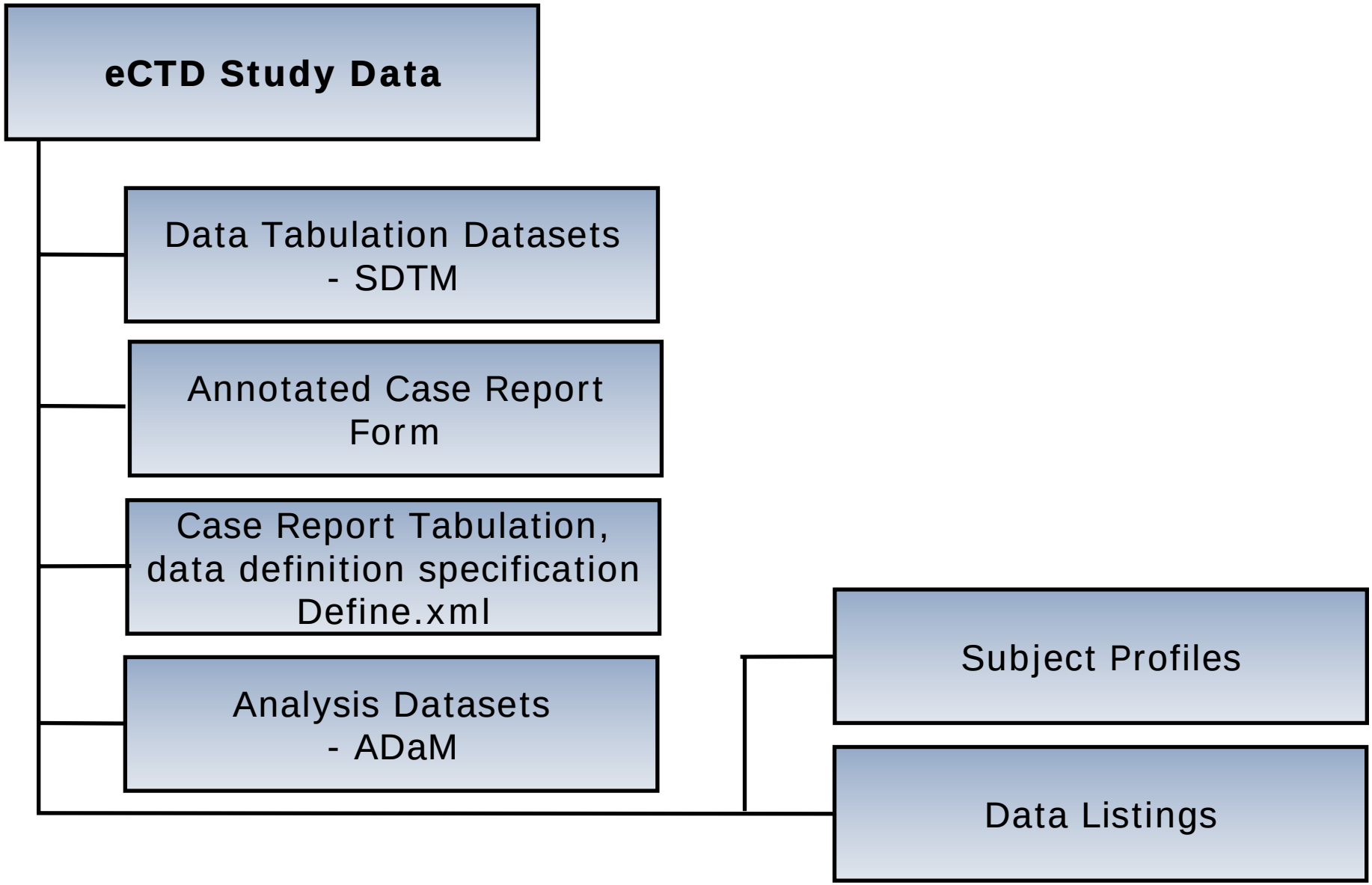
Providing Regulatory Submissions in Electronic Format — Human Pharmaceutical Product Applications and Related Submissions Using the eCTD Specifications

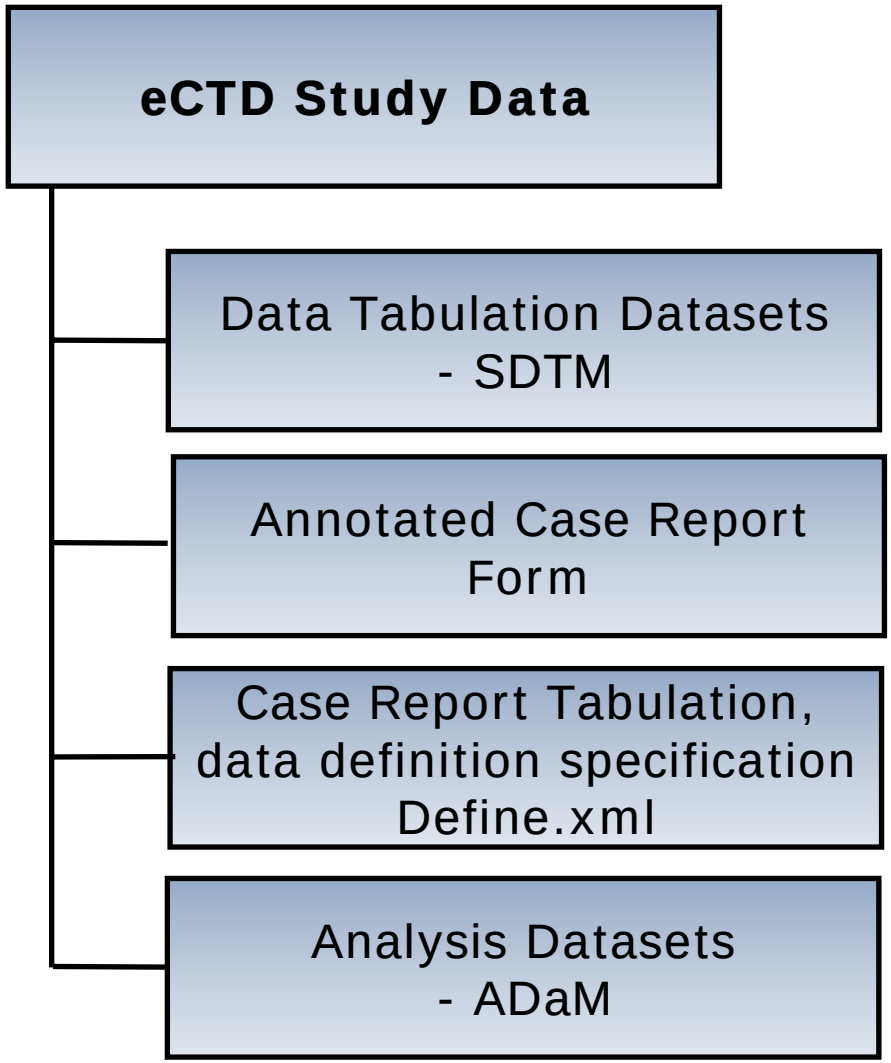
U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

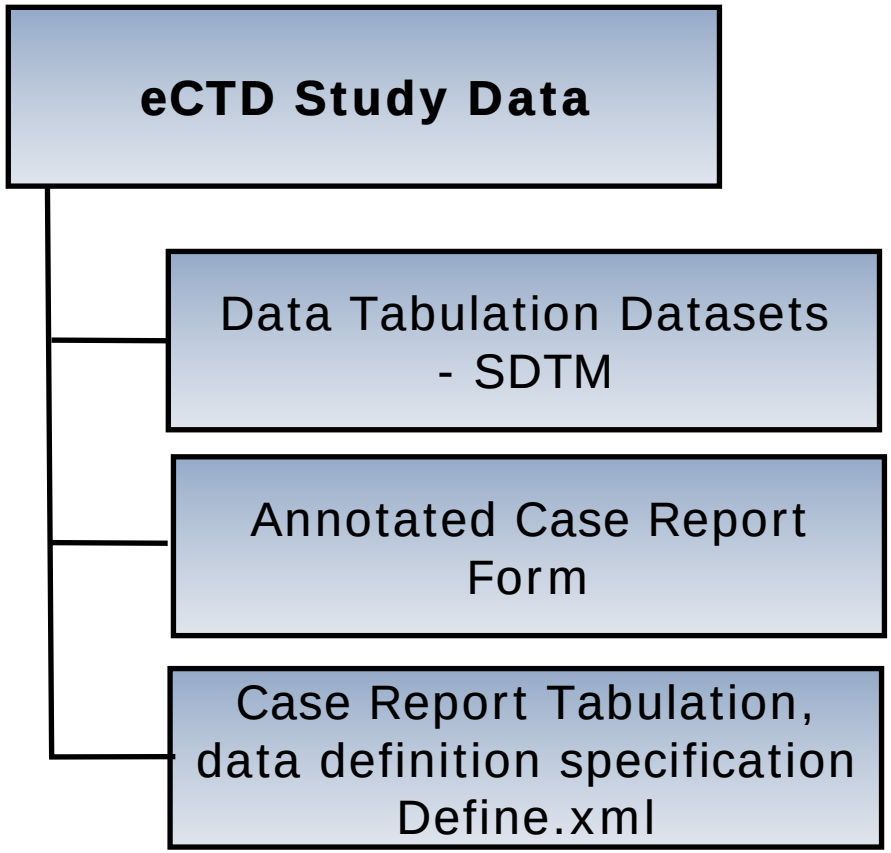
April 2006
Electronic Submissions

Revision 1

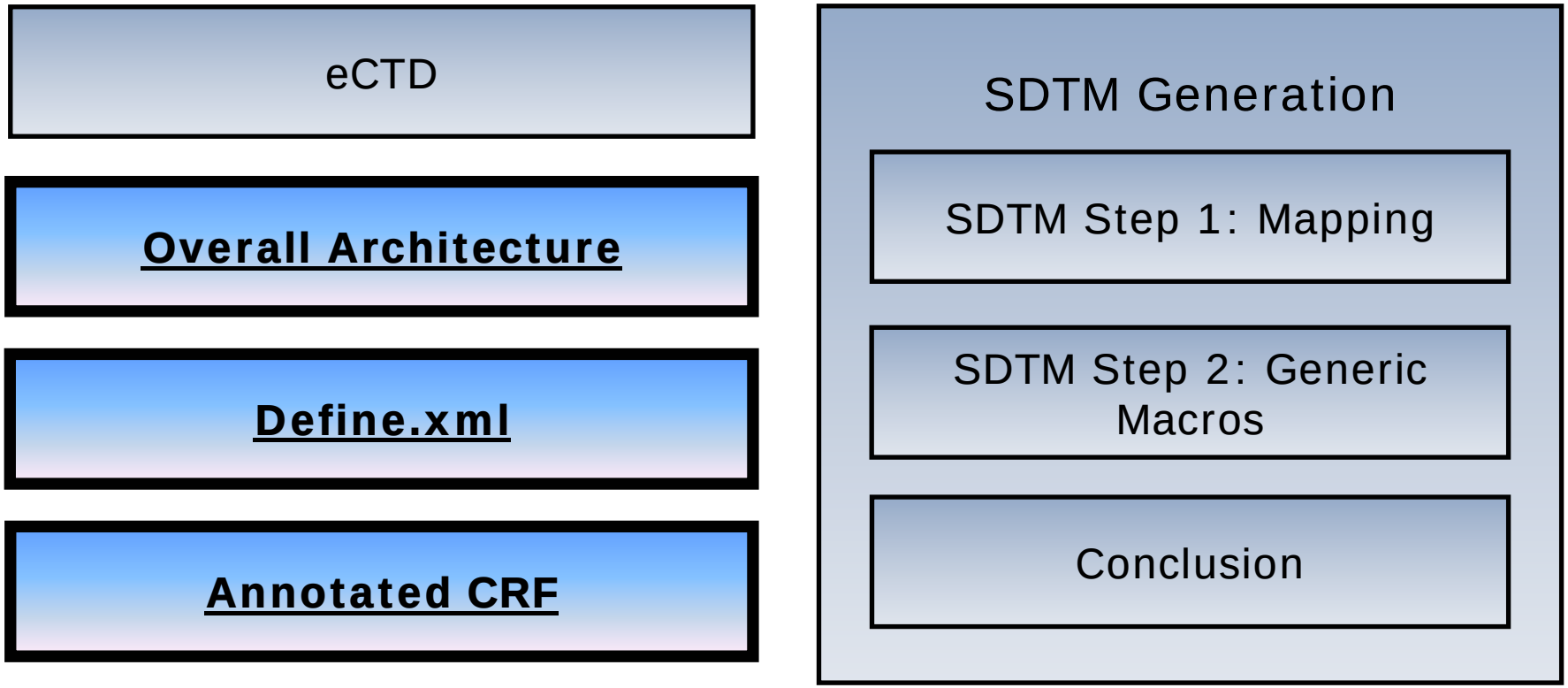
The study data specification


eCTD Study Data
Data Tabulation Datasets - SDTM
Annotated Case Report Form
Case Report Tabulation, data definition specification Define.xml
Analysis Datasets - ADaM
Subject Profiles
Data Listings

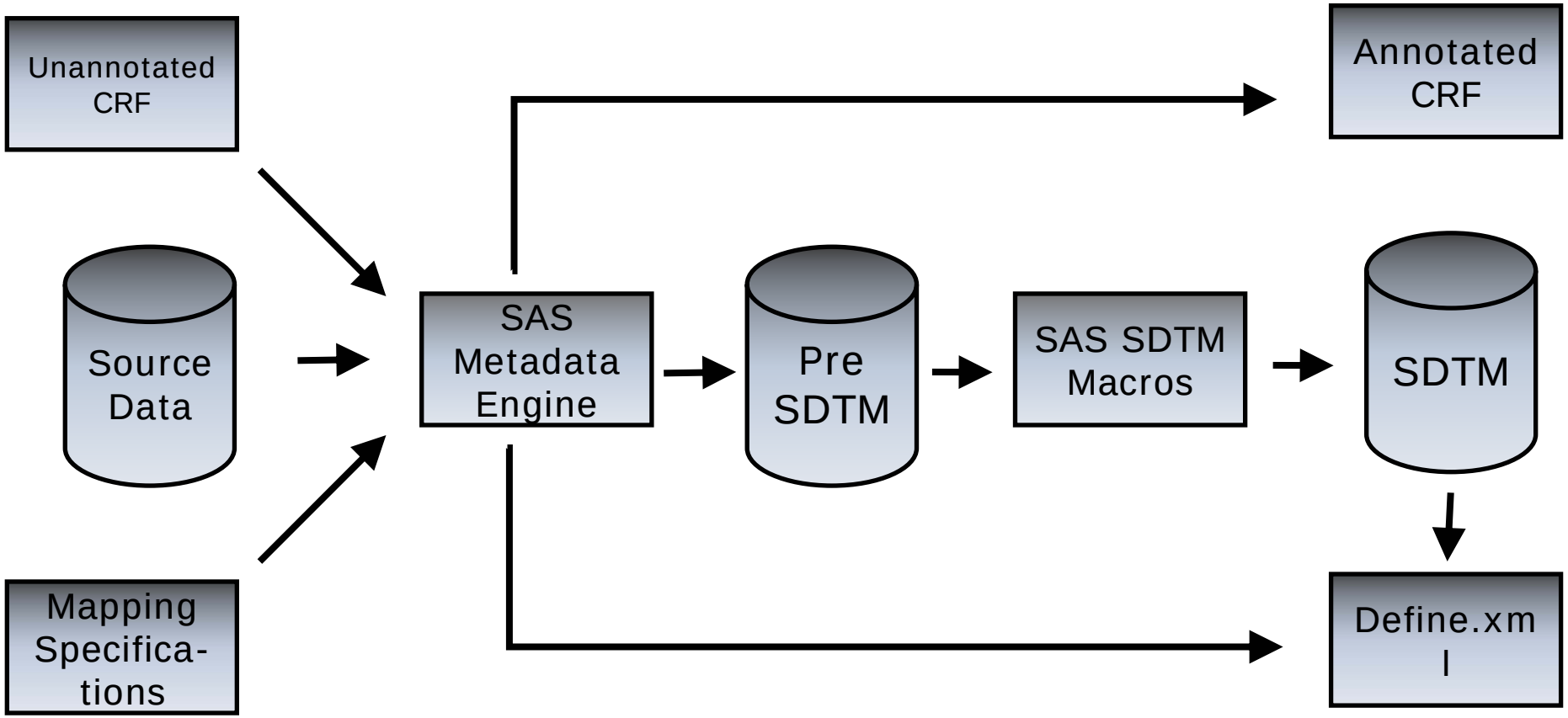

eCTD Study Data
Data Tabulation Datasets
- SDTM
Annotated Case Report
Form
Case Report Tabulation,
data definition specification
Define.xml
Analysis Datasets
- ADaM


eCTD Study Data
Data Tabulation Datasets - SDTM
Annotated Case Report Form
Case Report Tabulation, data definition specification Define.xml

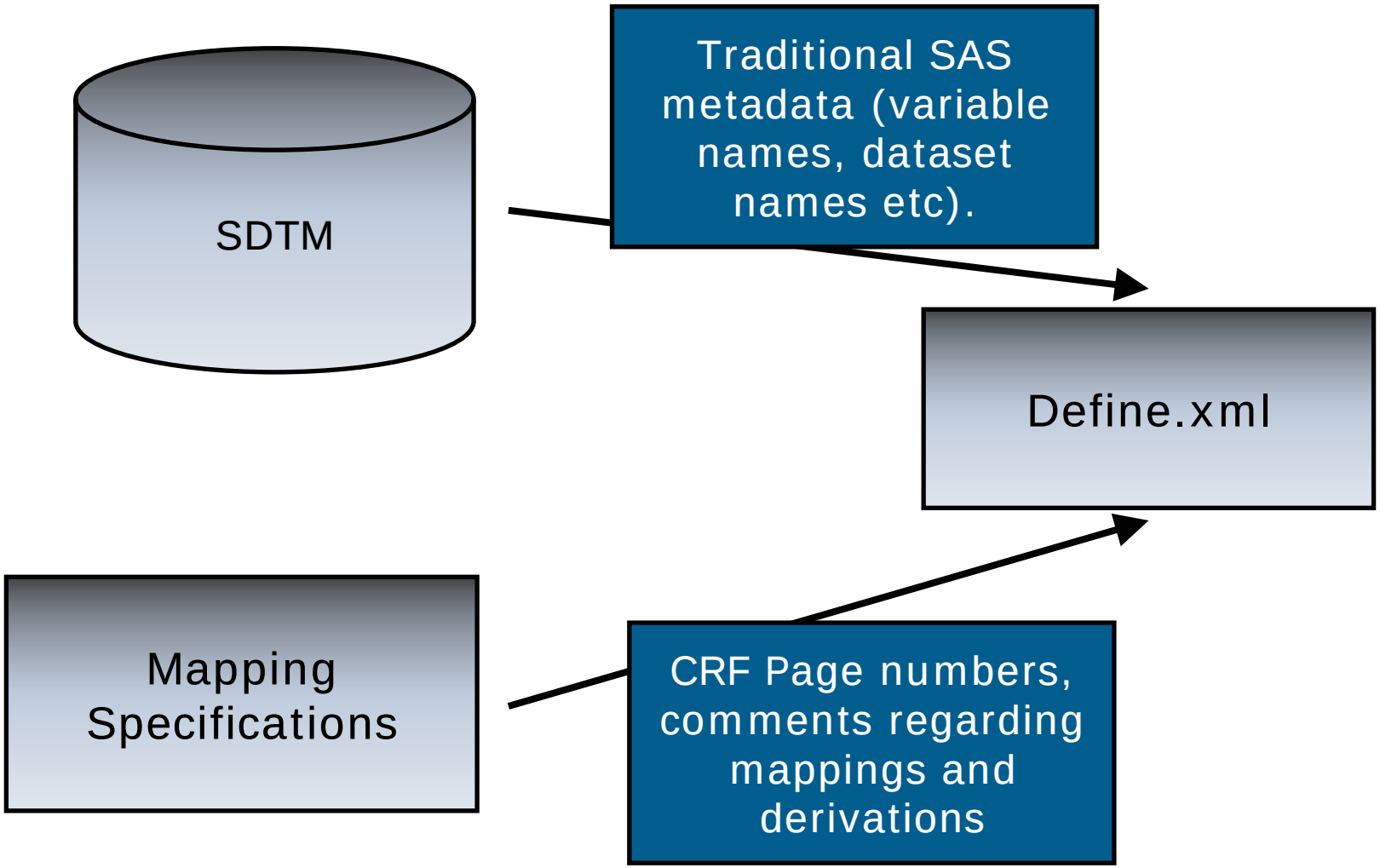
Presentation - Overview



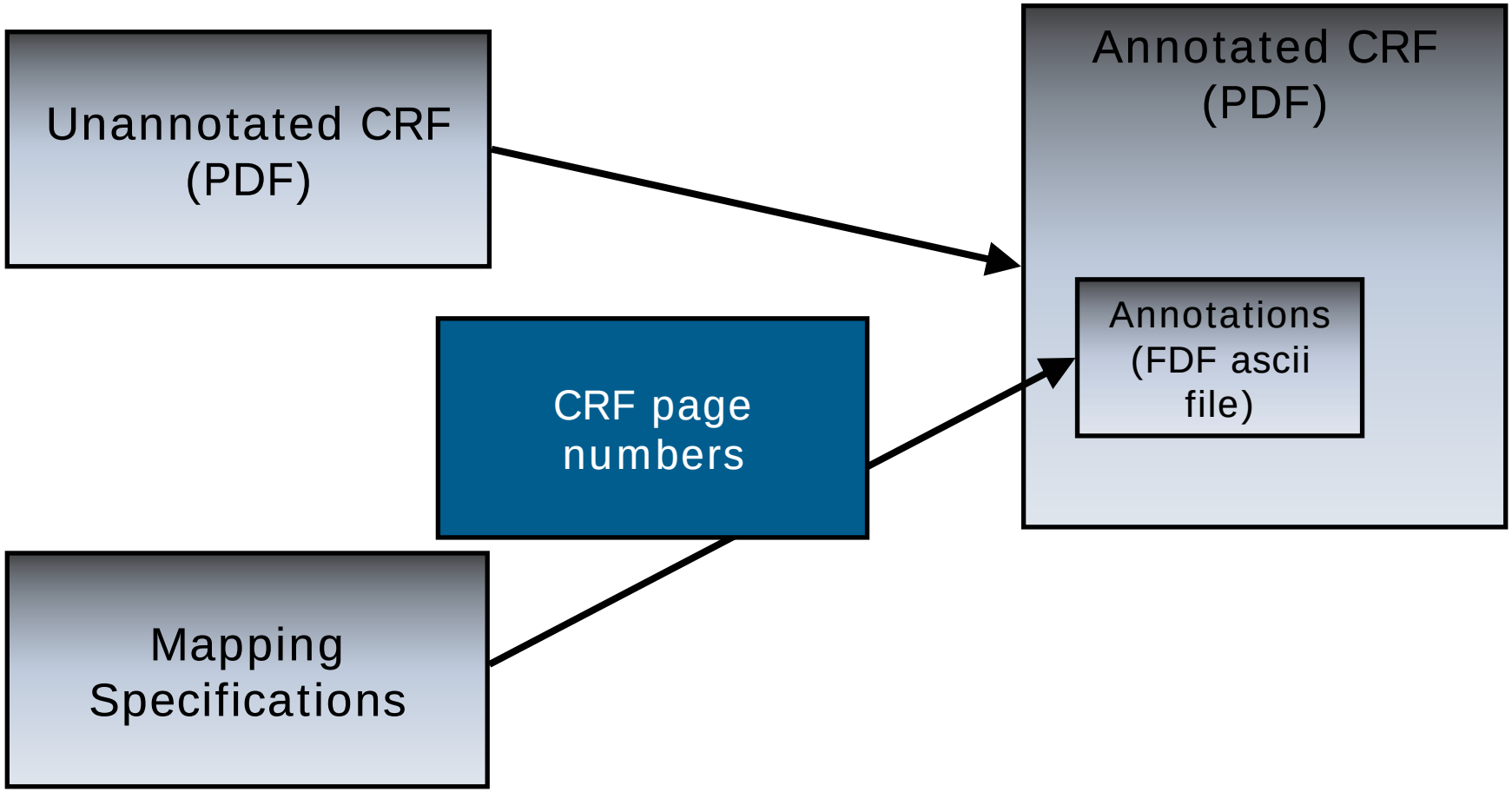
Overall Architecture



Create Define.xml

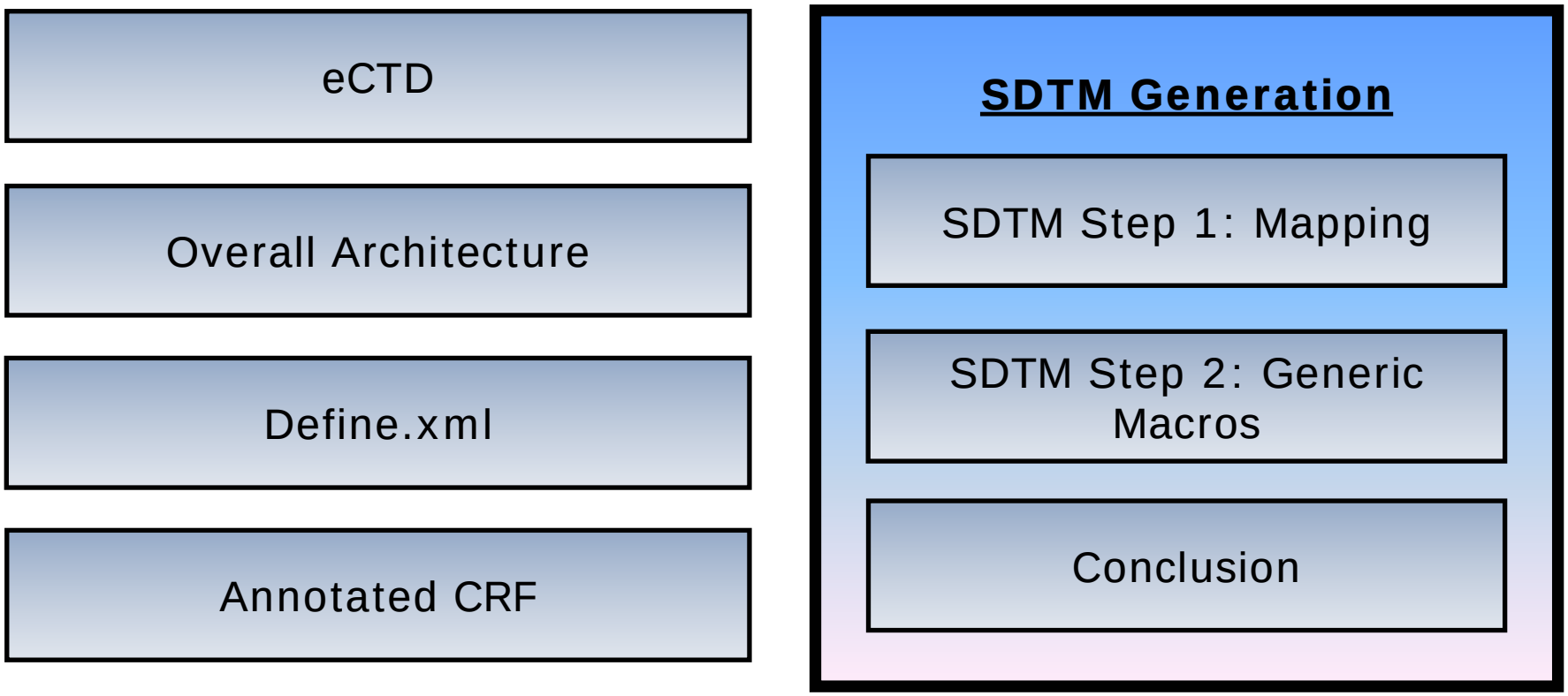


Annotate CRF with SAS

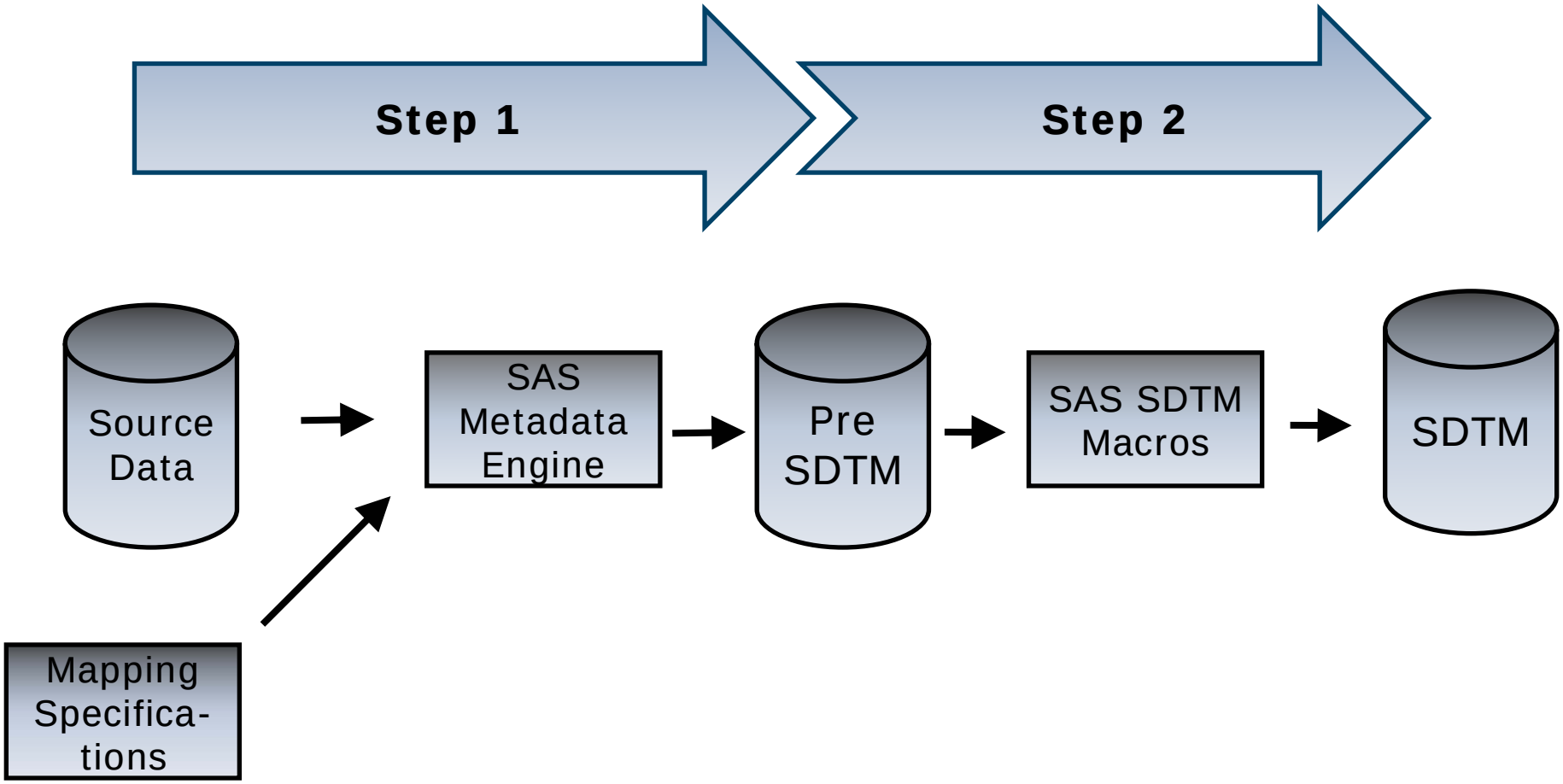


MORE INFORMATION: "*Using SAS to Speed up Annotating Case Report Forms in PDF Format prepared by Dirk Spruck and Monika Kawohl (PharmaSug 2004 paper CC02)*"

Presentation - Overview



SDTM Generation



SDTM Generation

- Use ETL tool ie SAS (Clinical) Data Integration Studio
- Use traditional SAS code
- Use spreadsheet with mappings entered to auto-generate the transforming SAS program

SDTM Generation

Macro name		Macro name
%Usubjid		%Relrec
%Domain		%Studyid
%Dtc		%Bl
%Dy		%Stc
%Seq		%Res
%LongChars		%SDTMify
%SuppqualCo		%PhysicalTransform

SDTM Generation – Macros (1/14)

%Usubjid

Creates the USUBJID variable (Unique subject ID) as a concatenation of SUBJID and Study ID

SDTM Generation – Macros (2/14)

%Domain

Creates the variable DOMAIN for all Pre SDTM datasets

SDTM Generation – Macros (3/14)

%DTC

Creates all --DTC variables in ISO8601 format based on date/time components stored as individual variables in Pre SDTM

SDTM Generation – Macros (4/14)

%Dy

Creates all --DY variables for each --DTC variable based on the RFSTDTC variable in the DM domain. The algorithm used is:

--DY = (date portion of --DTC) - (date portion of RFSTDTC) + 1 if --DTC is on or after RFSTDTC

--DY = (date portion of --DTC) - (date portion of RFSTDTC) if --DTC precedes RFSTDTC

SDTM Generation – Macros (5/14)

%Seq

Sorts datasets and assigns --SEQ variables as a unique sequence number within each USUBJID

SDTM Generation – Macros (6/14)

%LongChars

Identifies text strings longer than 200 characters and splits them up into multiple variables

SDTM Generation – Macros (7/14)

%SuppqualCo

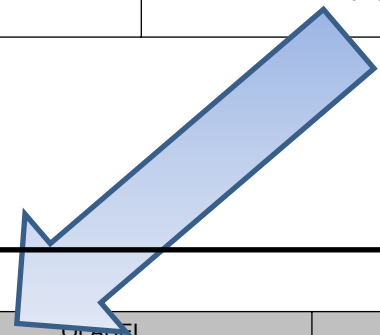
Creates supplementary datasets from extra columns on Pre SDTM datasets

Also creates CO datasets where relevant

SDTM Generation – Macros (7/14)

AE (Pre SDTM)

STUDYID	DOMAIN	USUBJID	AESEQ	AETERM	AESER	…	SQ_AESOTHC
SCUB01	AE	123	1	TRAUMA	Y	…	Neck Injury



SUPPAE

STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL
SCUB01	AE	123	AESEQ	1	AESOTHC	Desc of Other Event	Neck Injury

AE (SDTM)

STUDYID	DOMAIN	USUBJID	AESEQ	AETERM	AESER	…
SCUB01	AE	123	1	TRAUMA	Y	…

SDTM Generation – Macros (8/14)

%Relrec

Creates RELREC dataset describing relationships between records in Pre SDTM datasets

SDTM Generation – Macros (8/14)

Relrec example:

CM number 11&12 is related to AE number 3

RELREC

STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	RELTYPE	RELID
SCUB01	AE	123	AESEQ	3		AECM1
SCUB01	CM	123	CMSEQ	11		AECM1
SCUB01	CM	123	CMSEQ	12		AECM1

SDTM Generation – Macros (8/14)

Relrec example:

AE number 3 is related to CM number 11&12

Pre SDTM CM dataset:

STUDYID	DOMAIN	USUBJID	CMSEQ	RELREC_AE	...
SCUB01	CM	123	11	AESPID=3	...
SCUB01	CM	123	12	AESPID=3	...
...	...	...	...	...	...

The macro identifies the variables starting with 'RELREC-', then finds the associated records in the AE datasets, and based on that creates the 3 RELREC records

SDTM Generation – Macros (9/14)

%Studyid

Creates the variable STUDYID with the relevant study ID based on what is assigned in the Pre SDTM DM domain

SDTM Generation – Macros (10/14)

%BI

Creates the --BL variables, based on a defined algorithm (for example, by flagging the last assessement prior to RFSTDTC as the baseline value)

Note:

The existence of this statistically derived variable in the SDTM model is under review by the CDISC teams and the variable is expected to disappear from the SDTM model at some stage. Until that happens it is suggested to use a crude method like the above to set it and to note in the define document that it will not in all cases match the advanced baseline flags in the ADaM datasets.

SDTM Generation – Macros (11/14)

%Stc

Creates the various results variables in the finding domains, based on the Pre SDTM variables --ORRES and --ORESSU (original result and unit)

SDTM Generation – Macros (12/14)

%Res

Checks if the value of --STRESC is in fact numeric. In that case it is transferred to --STRESN

SDTM Generation – Macros (13/14)

%SDTMify

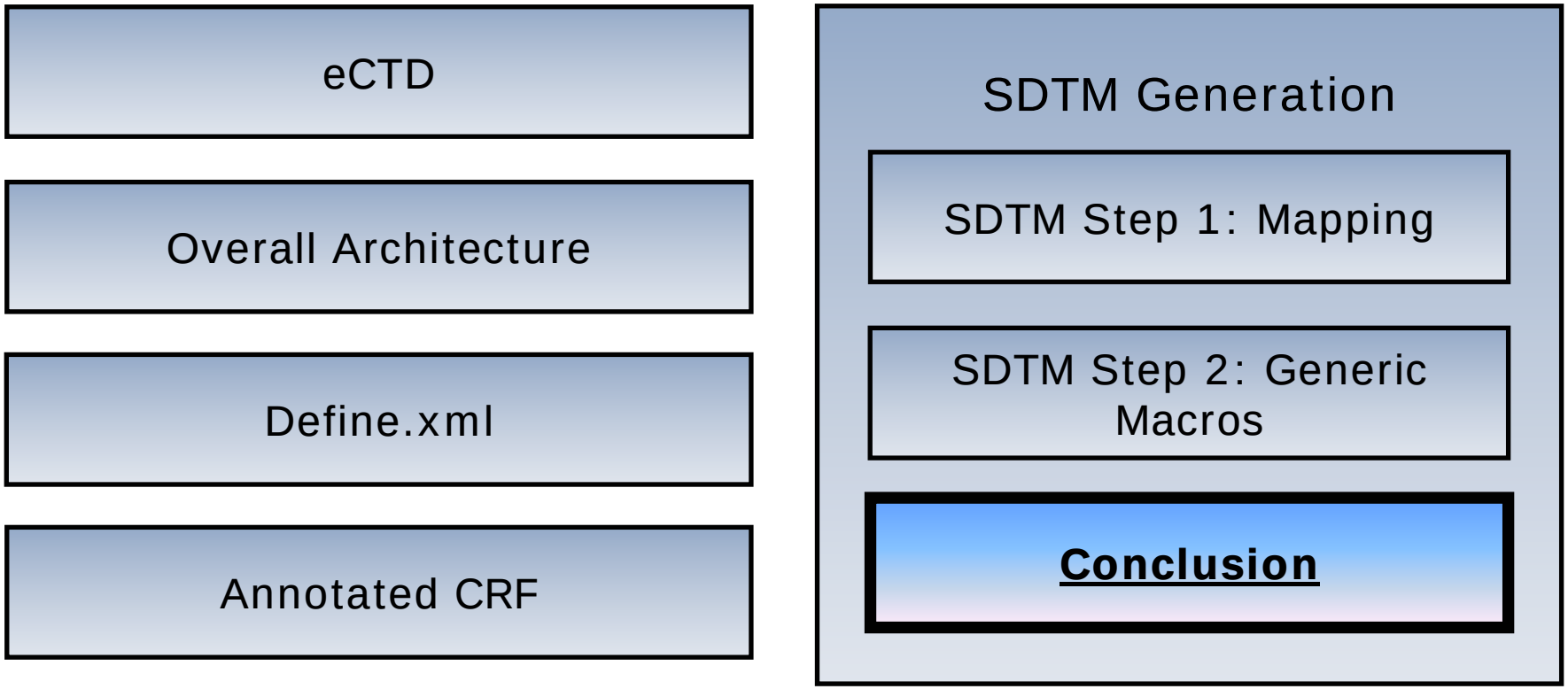
Assigns labels and exact formats to all the variables in the created datasets and sorts dataset records and columns according to agreed rules

SDTM Generation – Macros (14/14)

%PhysicalTransform

Turns the created SAS datasets into SAS transport files

Presentation - Overview



SDTM Generation