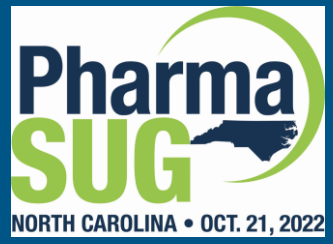
Gustav Bernard
Gustav Bernard is an Associate Director at IQVIA who has been with the company for 18 years.
His work focuses on the implementation of CDISC Standards (SDTM, ADaM and Define-XML) within IQVIA.
He is currently Leading ADaM Innovation @iqvia.

He has also created the Define-XML 2.0 automation process within IQVIA, the ADaM Spec Generator application and the ADaM Designer application.

Gustav earned a Bachelor of Business in computer science from the University of the Orange Free State in South Africa in 2004.



ADaM Automation Roadblocks

NC PharmaSUG 2022

Gustav Bernard, IQVIA, Associate Director

© 2021. All rights reserved. IQVIA® is a registered trademark of IQVIA Inc. in the United States, the European Union, and various other countries.

ADaM Automation Roadblocks

Making Sense of a Data Universe in Biostatistics

Digitize
Transforming ‘analog’ content into digital building blocks for creation and consumption

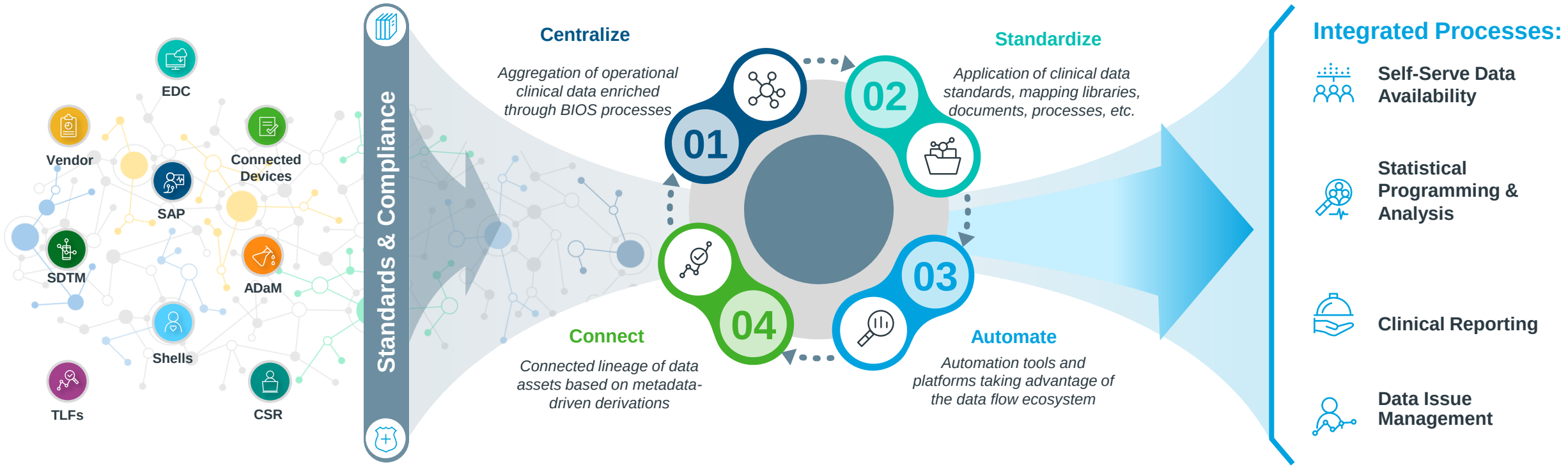
Standardize
Establish reusable, connected content with supporting processes to drive compliance

Centralize
Aggregating ‘data’ into a single organized, governed single source of truth

Automate
Identify opportunities to drive efficiency and quality across manual and repetitive tasks

Better BIOS Solutions

A robust BIOS solution that meets the new demands of the market



100% Data Ingestion Automation

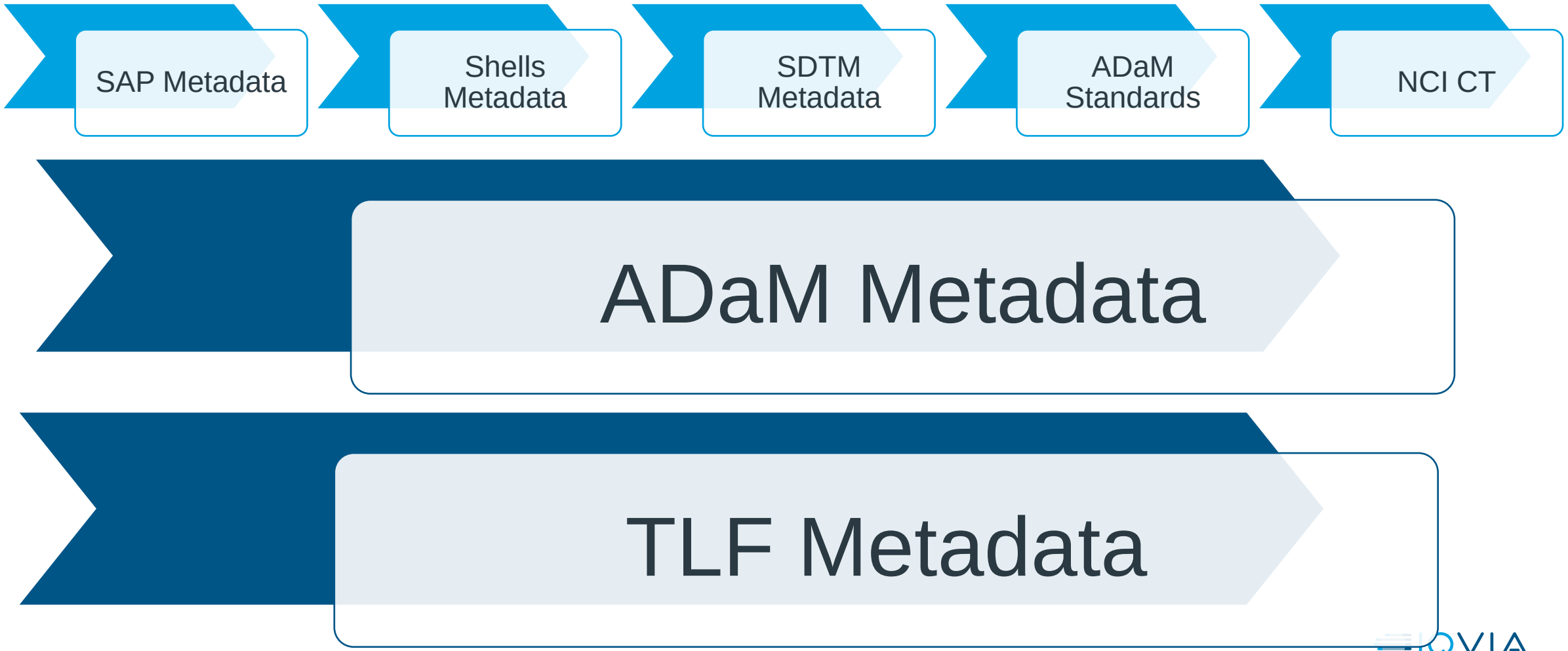
100% Data Consumption Automation

Real-Time Data Flow & Access

Data-Driven Standardized Process

Connecting Processes

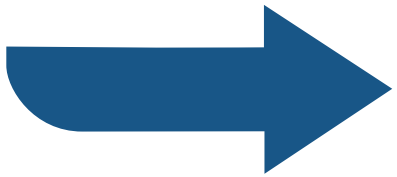
Web Based with Single backend Database (ADaM Focus)



Change Management

Continues Evolution of processes

- Excel Template
- Standard in PDF document



- VBA Application
- CDISC Standards Template with Customized Metadata
- NCI CT Terminology
- SDTM Metadata
  - Domain Level
  - Variable Level
  - VLM Level
  - Variable and VLM Terminology

- Web Based Version
- API Link To CDISC Library
- CDISC Standards Template with Customized Metadata
- API link To NCI CT Terminology
- SDTM Metadata
  - Domain Level
  - Variable Level
  - VLM Level
  - Variable and VLM Terminology

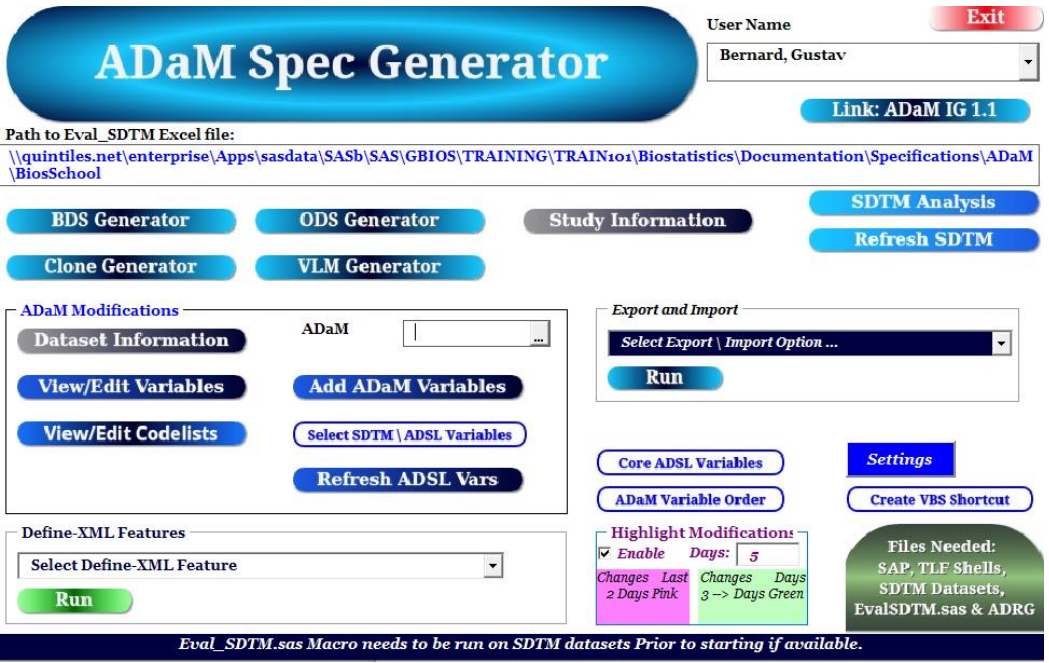
Change Management

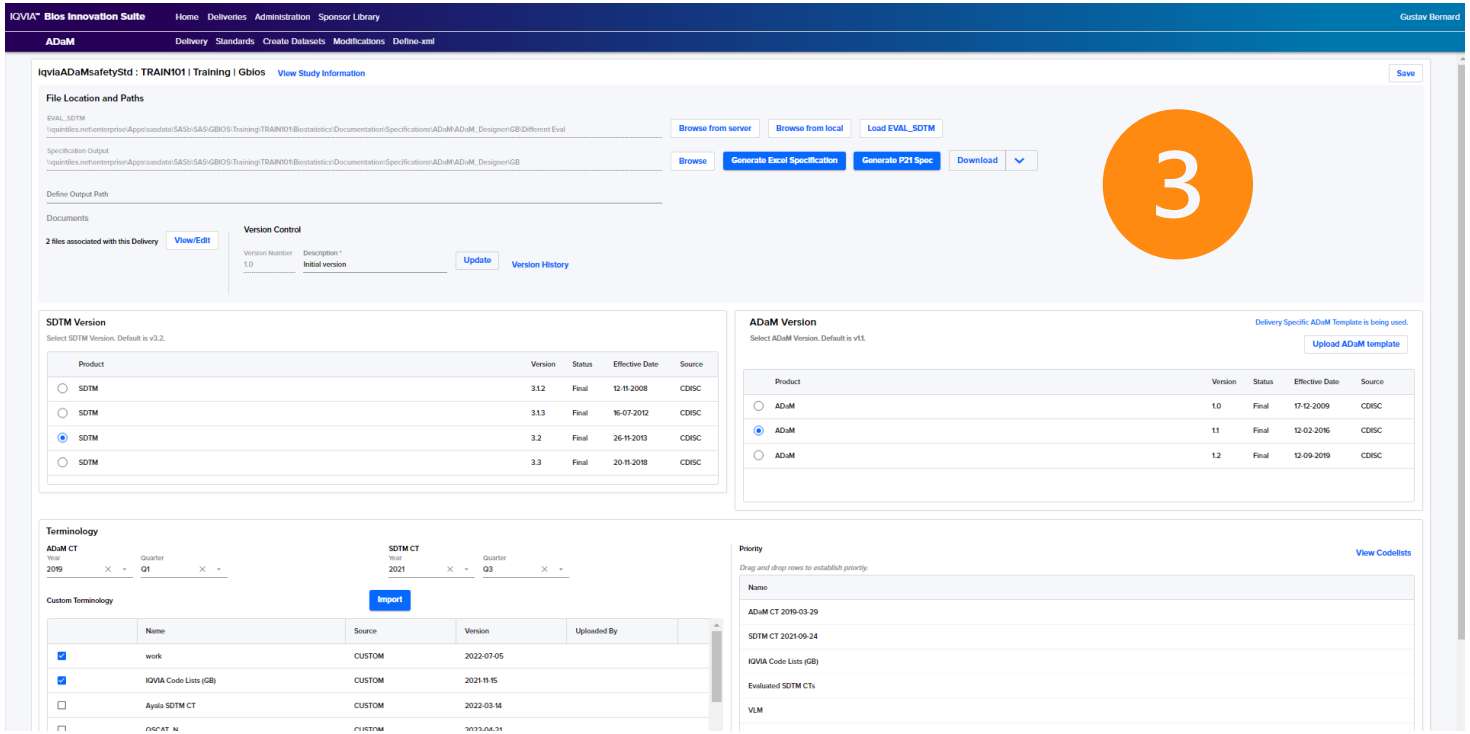
Continues Evolution of processes

1

Final Variable Order in Dataset #	Variable Name	Variable Label	Type	Length	Controlled Terms, Codelist or Format	Origin	Computational Algorithm or Method (CDISC only)	Details of derivation	Keep (Y/N)
Standard SDSL Variables									
1	STUDYID	Study Identifier	Char	20		SDTM.DM			Y
2	USUBJID	Unique Subject Identifier	Char	40		SDTM.DM			Y
3	SUBJID	Subject Identifier for the Study	Char	20		SDTM.DM			Y
4	SITEID	Study Site Identifier	Char	4		SDTM.DM			Y
6	BRTHDT	Date of Birth	Char	19	ISO8601	SDTM.DM		date part of DM.BRTHDTC	Y
7	BRTHDTN	Date of Birth (N)	Num	8	DATE9.	Derived		convert BRTHDT to a numeric value	Y
8	AGE	Age	Num	8		SDTM.DM			Y
9	AGEU	Age Units	Char	10	AGEU	SDTM.DM			Y
10	SEX	Sex	Char	2	SEX	SDTM.DM			Y
11	SEXN	Sex (N)	Num	8	SEXN	Derived			Y

2

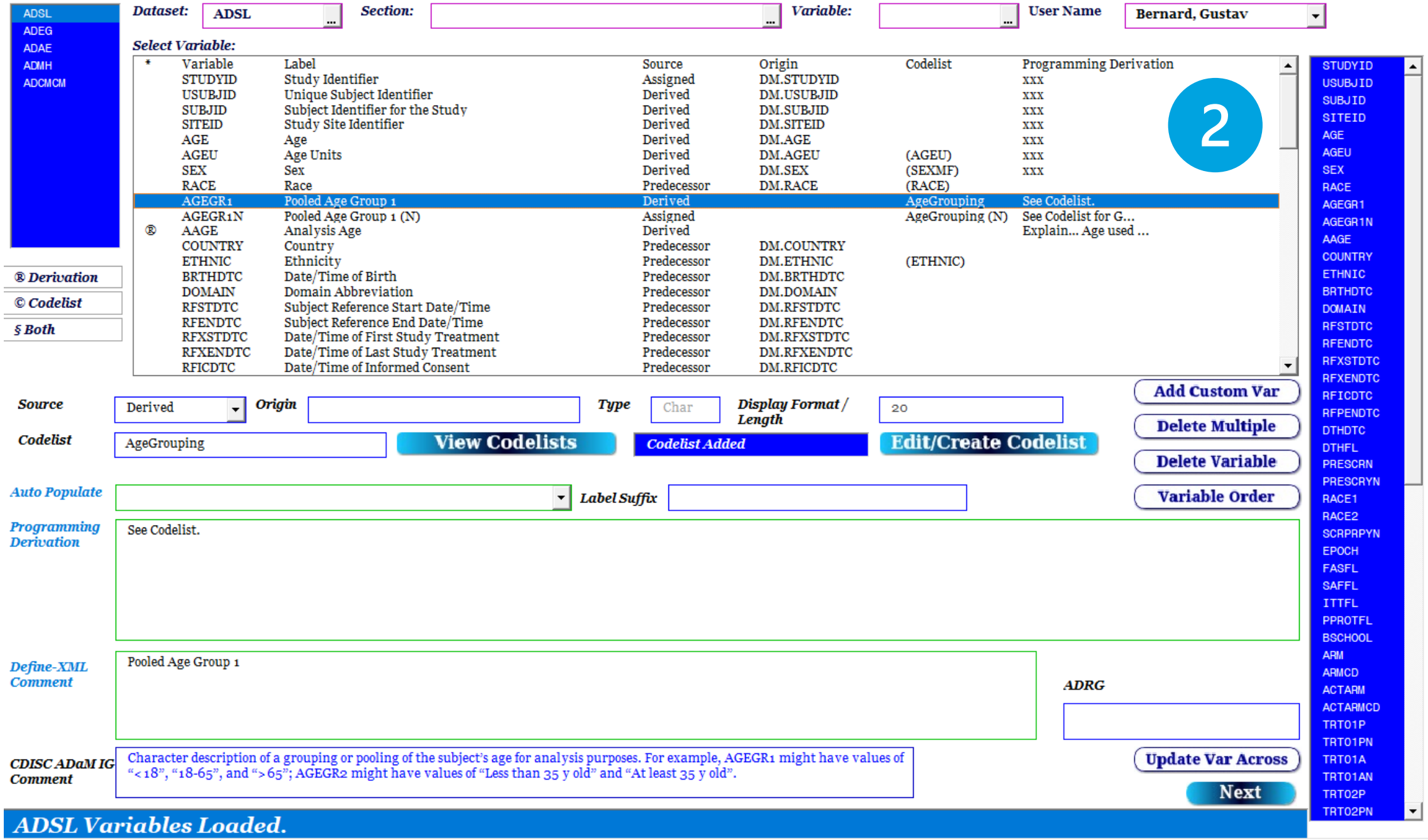


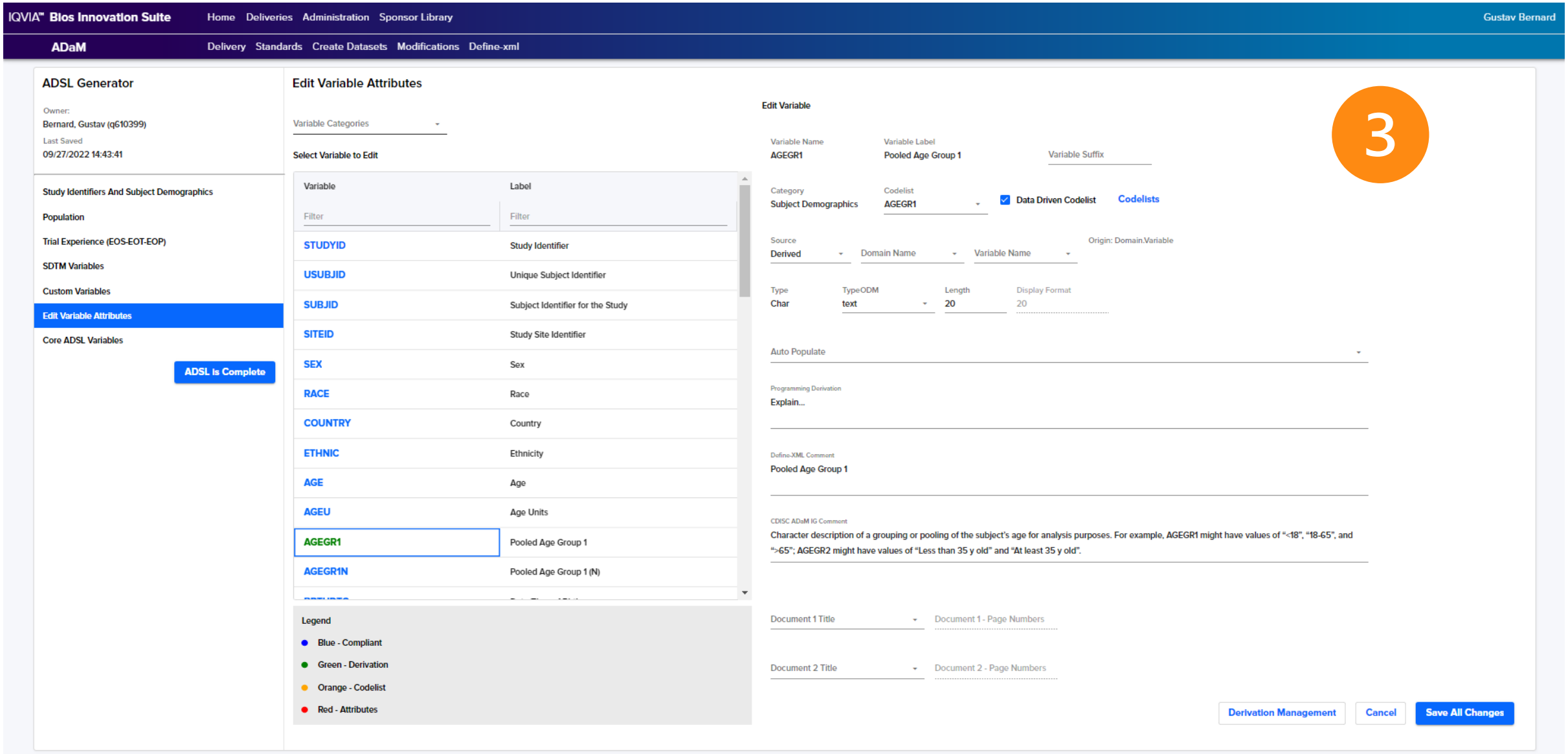


3

Traditional Spec

Final Variable Order in Dataset #	Variable Name	Variable Label	Type	Length	Controlled Terms, Codelist or Format	Origin	Computational Algorithm or Method (CDISC only)	Details of derivation	K (
Standard SDSL Variables									
1	STUDYID	Study Identifier	Char	20		SDTM.DM			Y
2	USUBJID	Unique Subject Identifier	Char	40		SDTM.DM			Y
3	SUBJID	Subject Identifier for the Study	Char	20		SDTM.DM			Y
4	SITEID	Study Site Identifier	Char	4		SDTM.DM			Y
6	BRTHDT	Date of Birth	Char	19	ISO8601	SDTM.DM		date part of DM.BRTHDTC	Y
7	BRTHDTN	Date of Birth (N)	Num	8	DATE9.	Derived		convert BRTHDT to a numeric value	Y
8	AGE	Age	Num	8		SDTM.DM			Y
9	AGEU	Age Units	Char	10	AGEU	SDTM.DM			Y
10	SEX	Sex	Char	2	SEX	SDTM.DM			Y
11	SEXN	Sex (N)	Num	8	SEXN	Derived			Y
	GWCP	Childbearing Potential	Char			SDTM.SUPPDM		SUPPDM.QVAL where SUPPDM.QNAM = "DMGWCPT"	
	GWCPN	Childbearing Potential (N)	Num	8					
12	RACE	Race	Char	50	RACE	SDTM.DM			Y
13	RACEN	Race (N)	Num	8	RACEN	Derived			Y
	RACEOTH	Race Other - Specify	Char	100		SDTM.SUPPDM		SUPPDM.QVAL where SUPPDM.QNAM ='DMRACEOT'	Y
	RACEAB	Race Abbreviated	Char	4	RACEAB	Derived			Y
14	ETHNIC	Ethnicity	Char	22	ETHNIC	SDTM.DM			Y
15	ETHNICN	Ethnicity (N)	Num	8	ETHNICN	Derived			Y
	ETHNICAB	Ethnic Abbreviated	Char	4	ETHNICAB	Derived			Y
30	ICDT	Date of Informed Consent	Num	8	DATE9.	Derived		Informed Consent date from SDTM.DM	Y

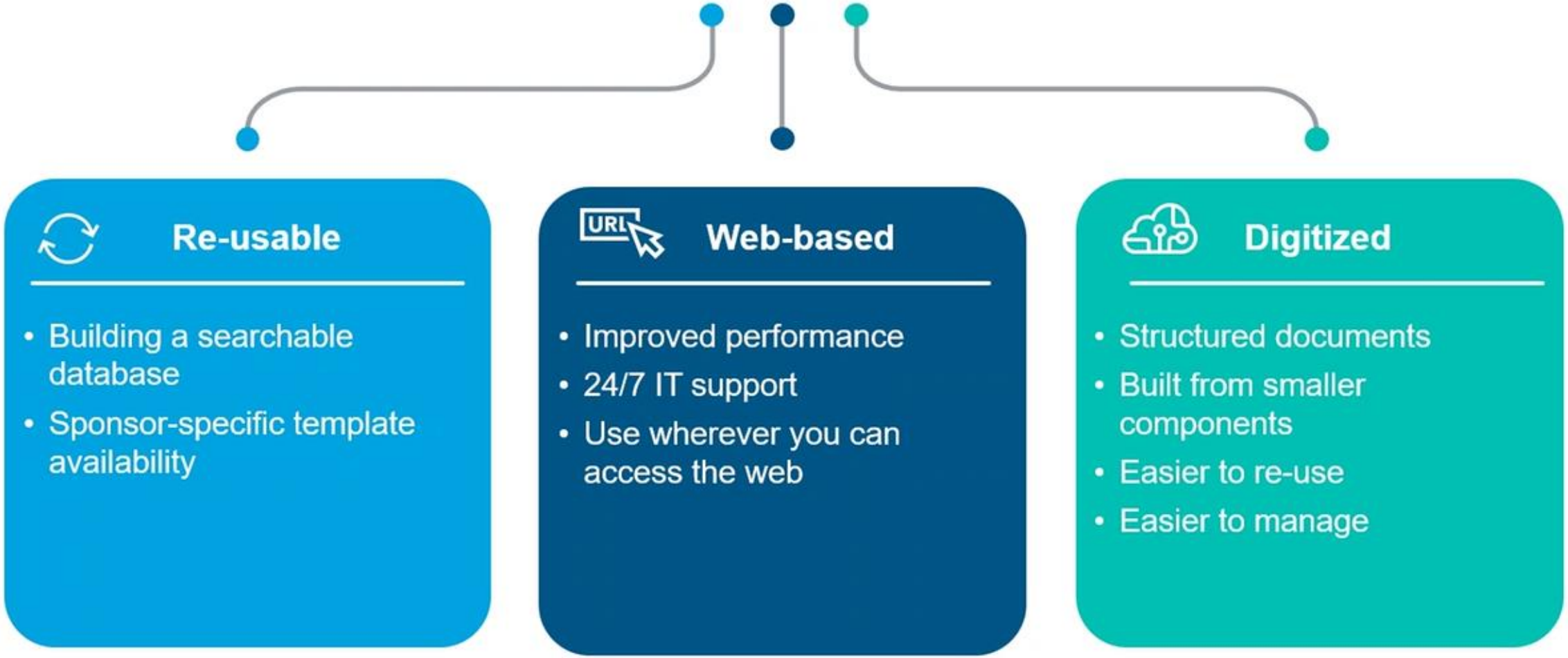

ADSL
ADEG
ADAE
ADMH
ADCMCM
Dataset: ADSL
Section:
Variable:
User Name Bernard, Gustav
Select Variable:
* Variable Label Source Origin Codelist Programming Derivation
STUDYID Study Identifier Assigned DM.STUDYID xxx
USUBJID Unique Subject Identifier Derived DM.USUBJID xxx
SUBJID Subject Identifier for the Study Derived DM.SUBJID xxx
SITEID Study Site Identifier Derived DM.SITEID xxx
AGE Age Derived DM.AGE xxx
AGEU Age Units Derived DM.AGEU (AGEU) xxx
SEX Sex Derived DM.SEX (SEXMF) xxx
RACE Race Predecessor DM.RACE (RACE)
AGEGR1 Pooled Age Group 1 Derived AgeGrouping See Codelist.
AGEGR1N Pooled Age Group 1 (N) Assigned AgeGrouping (N) See Codelist for G...
® AAGE Analysis Age Derived Explain... Age used ...
COUNTRY Country Predecessor DM.COUNTRY
ETHNIC Ethnicity Predecessor DM.ETHNIC (ETHNIC)
BRTHDTC Date/Time of Birth Predecessor DM.BRTHDTC
DOMAIN Domain Abbreviation Predecessor DM.DOMAIN
RFSTDTC Subject Reference Start Date/Time Predecessor DM.RFSTDTC
RFENDTC Subject Reference End Date/Time Predecessor DM.RFENDTC
RFXSTDTC Date/Time of First Study Treatment Predecessor DM.RFXSTDTC
RFXENDTC Date/Time of Last Study Treatment Predecessor DM.RFXENDTC
RFICDTC Date/Time of Informed Consent Predecessor DM.RFICDTC
2
STUDYID USUBJID SUBJID SITEID AGE AGEU SEX RACE AGEGR1 AGEGR1N AAGE COUNTRY ETHNIC BRTHDTC DOMAIN RFSTDTC RFENDTC RFXSTDTC RFXENDTC RFICDTC RFPENDTC DTHDTC DTHFL PRESCRN PRESCRYN RACE1 RACE2 SCRPRPYN EPOCH FASFL SAFFL ITTFL PPROTFL BSCHOOL ARM ARMCD ACTARM ACTARMCD TRT01P TRT01PN TRT01A TRT01AN TRT02P TRT02PN
® Derivation
© Codelist
§ Both
Source Derived
Origin
Type Char
Display Format / Length 20
Codelist AgeGrouping
View Codelists
Codelist Added
Edit/Create Codelist
Add Custom Var
Delete Multiple
Delete Variable
Variable Order
Auto Populate
Label Suffix
Programming Derivation
See Codelist.
Define-XML Comment
Pooled Age Group 1
ADRG
CDISC ADaM IG Comment
Character description of a grouping or pooling of the subject's age for analysis purposes. For example, AGEGR1 might have values of "< 18", "18-65", and "> 65"; AGEGR2 might have values of "Less than 35 y old" and "At least 35 y old".
Update Var Across
Next
ADSL Variables Loaded.



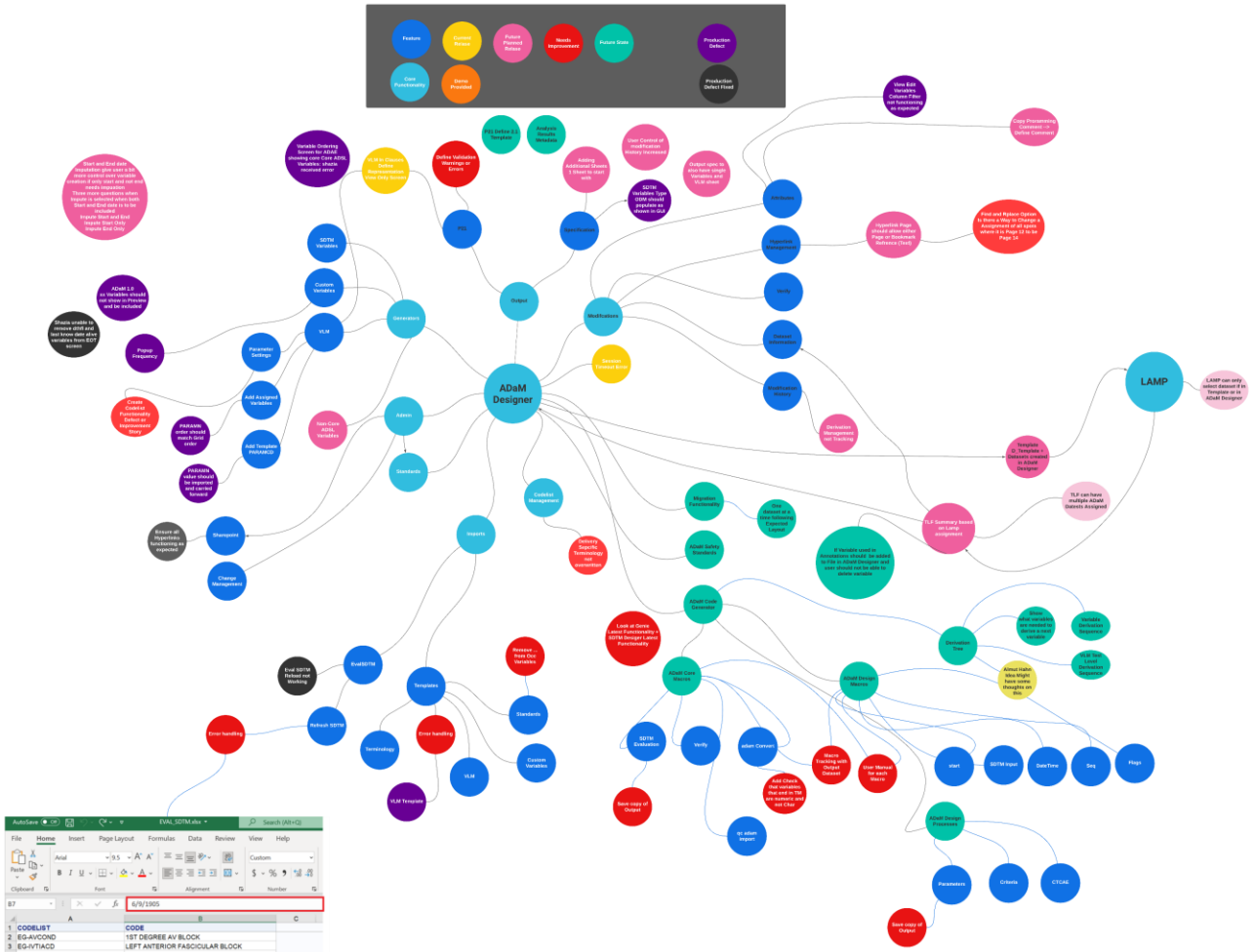
PharmaSUG
NORTH CAROLINA • OCT. 21, 2022

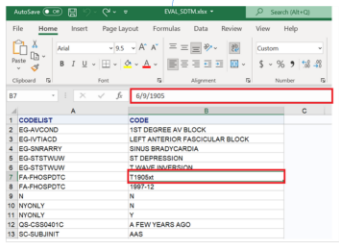
IQVIA

Why have a Tool \ Application

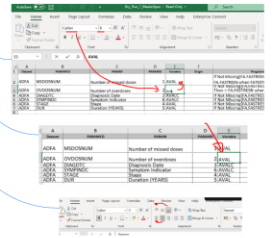


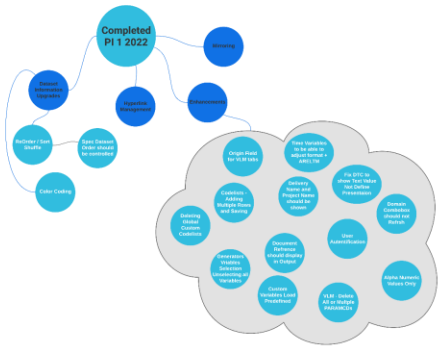
ADaM Automation Keeping Track





CODELIST	CODE
EQ-AVCOND	1ST DEGREE AV BLOCK
EQ-IVTIACD	LEFT ANTERIOR FASCICULAR BLOCK
EQ-SINRARRY	SINUS BRADYCARDIA
EQ-STSTWUW	ST DEPRESSION
EQ-STSTWUW	T WAVE INVERSION
FA-FHOSPDTC	T1905st
FA-FHOSPDTC	1997-12
N	N
NYONLY	N
NYONLY	Y
QS-CSS0401C	A FEW YEARS AGO
SC-SUBJINIT	AAS

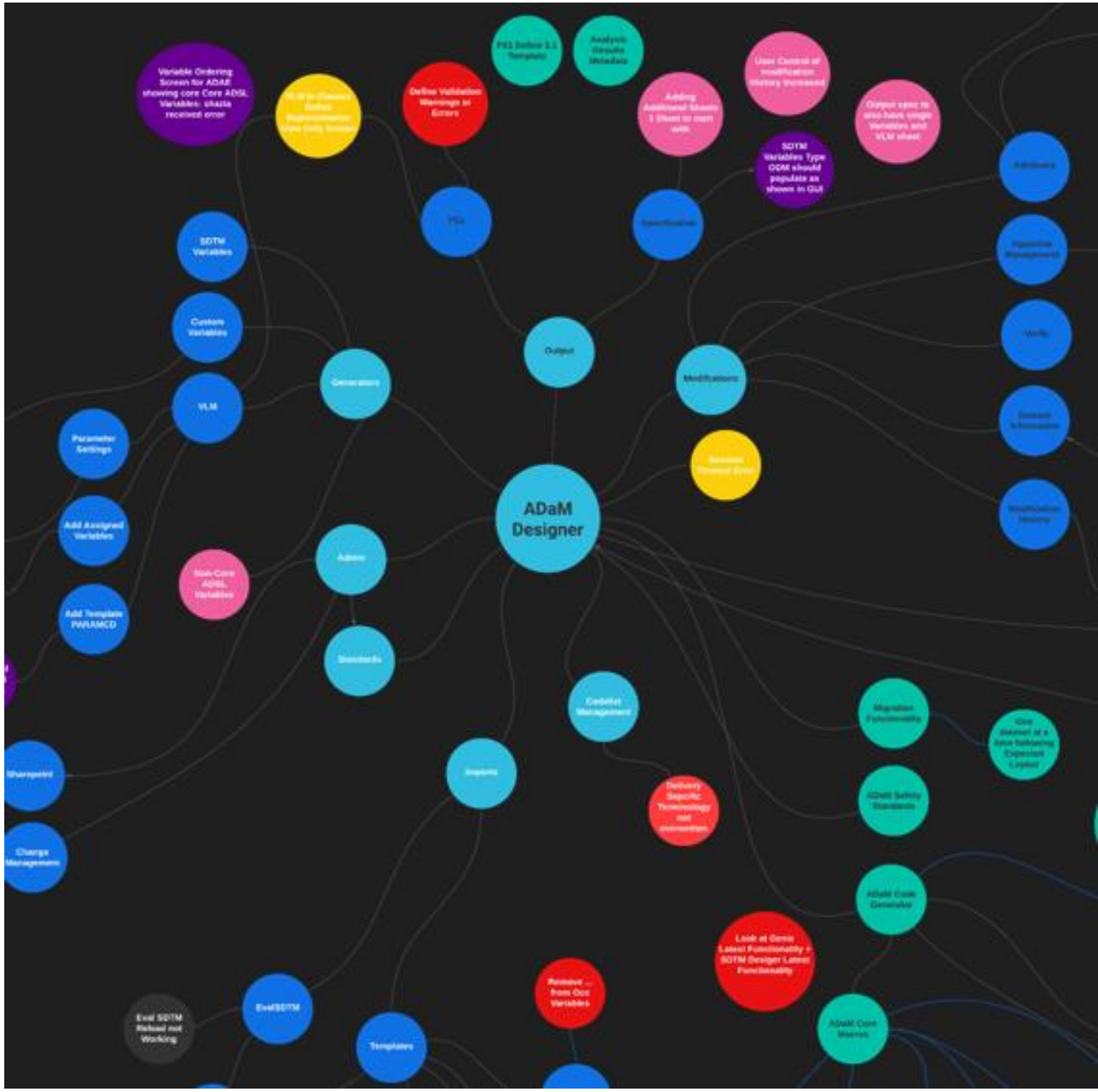




PharmaSUG
NORTH CAROLINA • OCT. 21, 2022

ADaM Automation Keeping Track

- Keeping Track
- What has been created
- What is being developed in current release
- What to develop in next release
- What to develop for the following release
- What functionality needs improvements
- What functionality is not working as expected
- What functionality needs to be investigated
- What linking with other applications needs to happen when



Automation Implementation

Tool vs Process vs Macro

- Tool

- Specification Metadata

- Should be setup to be able to handle SAS \ R \ Python

- Process

- For Example, Macro that Runs over all SDTM datasets to create SDTM Evaluation that feeds into Tool.

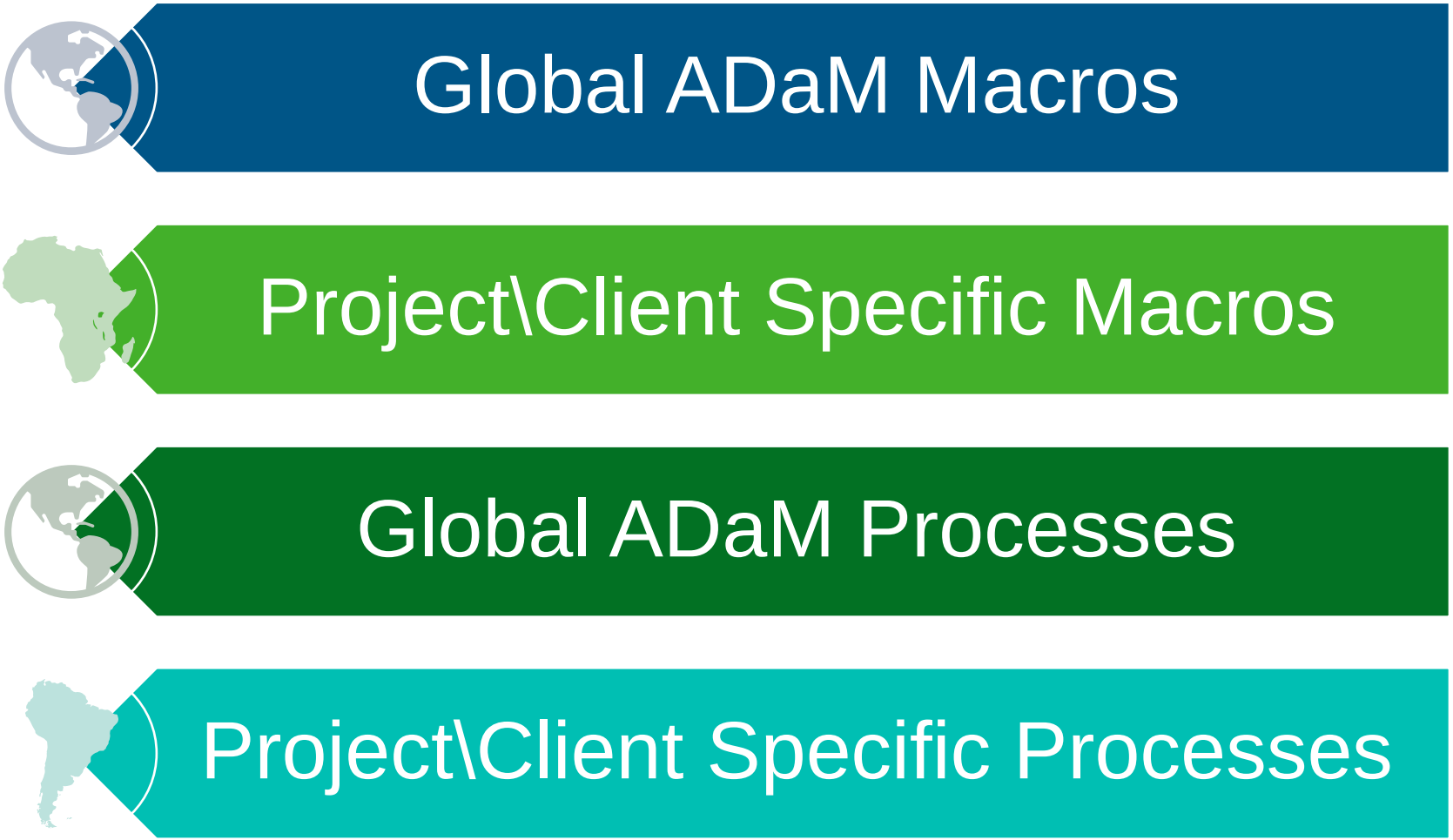
- Macro that runs over SDTM Evaluation and Create Excel Output with recommended assignment of Parameters from SDTM → ADaM
  - Output Read in by SAS programs to create Parameter variables
  - Output Read by tool to create VLM Metadata

- Macros

- Macro that uses Metadata from either Tools or Processes to create variables or add additional rows of data.

Innovation

Macro Management



ADaM Macros

- # of Parameters
- Name and Description for each Parameter
- Set allowed contents for parameter (Restricted vs not)
  - Restricted when Set of expected values allowed
    - › Add Restricted list of allowed Values (e.g., 0 or 1)
  - Free Text
  - Selection from Metadata \ Where Clause
  - Default Values
- Variable(s) to be derived using Macro
  - What variables should be auto assigned (All Predecessor variable handled by one macro)
- Expected Single call or Multiple call
  - Datetime macro that will create DTN, TM, DTM and DY variables only needs to called once.
- Expected Code Level / Should always be called
- Code Levels
  - Initialize
  - Customize
  - Formalize \ Standardize
  - Finalize
- Each macro should have adequate defensive programming for each macro parameter with clear messaging if parameter used incorrectly.
- Version Control of each Macro
- Creating software dependent macros that are controlled via parameter calls

Innovation

Code Management (Code Creation and Maintenance)

Within a controlled system?

- Code maintenance becomes a concern and can lead to rework.
- Highly controlled

Outside of a system?

- Tried and tested works but uncontrolled

Best of Both Worlds

- System Controls Certain code levels
  - Initialize
  - Formalize \ Standardize
  - Finalize
- Output each level per dataset
- User calls each Level to execute in program and can add Customized Level.

Innovation

Decision Tree

Variable Level

- Variable Derivation Sequence
- Variable Dependencies
- Automated until User takes control
- Code Level and Variable Priority

VLM Level

- Test Level Derivation Sequence
- Test Level Dependencies
- Automated until User takes control
- Code Level and Test Priority

Traditional vs Automated

Traditional (AOCC01FL)

Initial Creation

- **Explain** Derivation in Specification
- **Code** Derivation in Production Program
- **Code** Derivation in QC Program

Maintenance (Rework)

- **Update** Derivation in Specification
- **Update** Derivation in Production Program
- **Update** Derivation in QC Program

Automated (AOCC01FL)

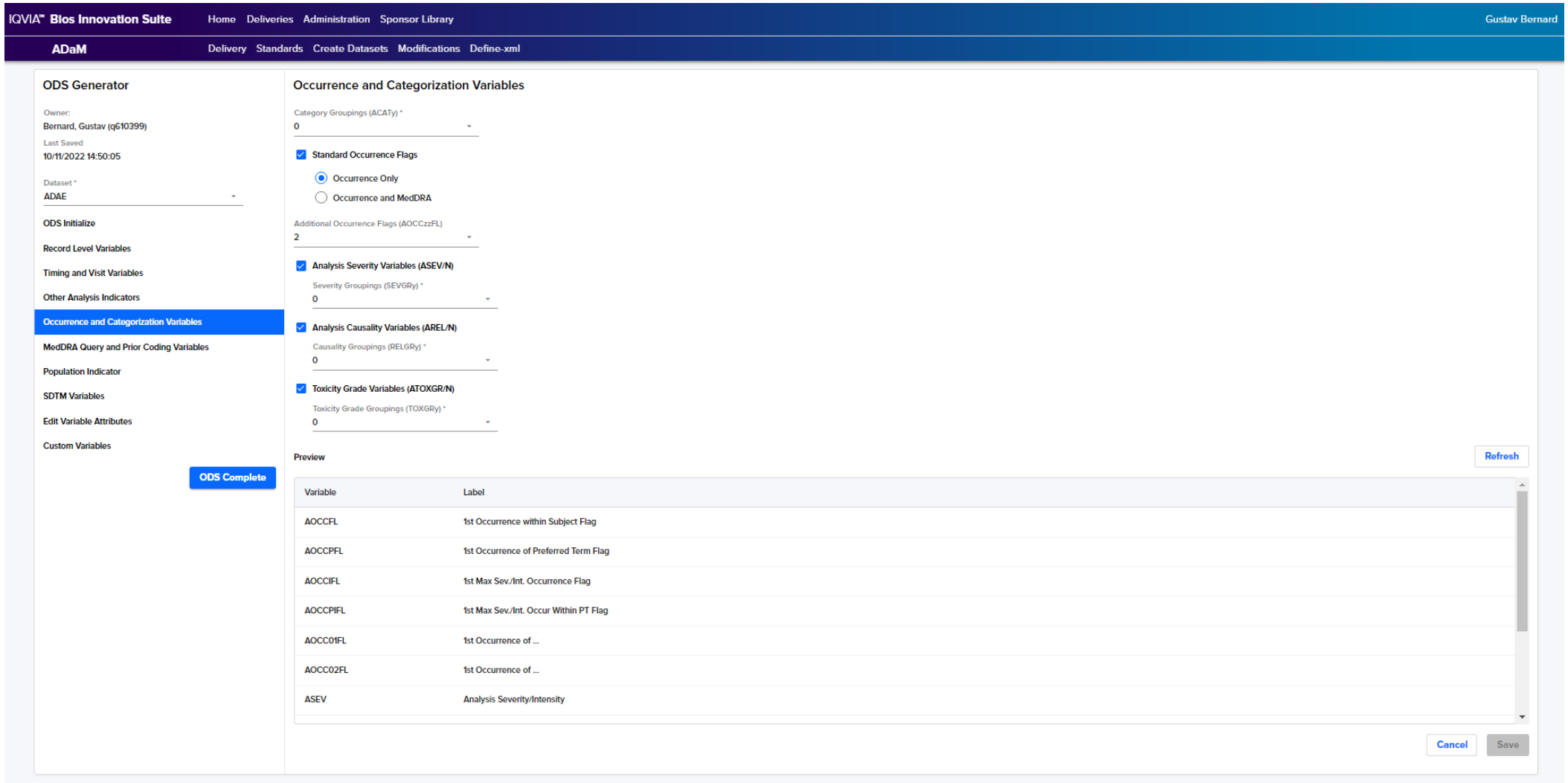
Initial Creation

- **Explain** Derivation in Specification
- Macro **Extracts** Derivation in Prod Prog
- QC Macro **Extracts** Derivation in QC Prog

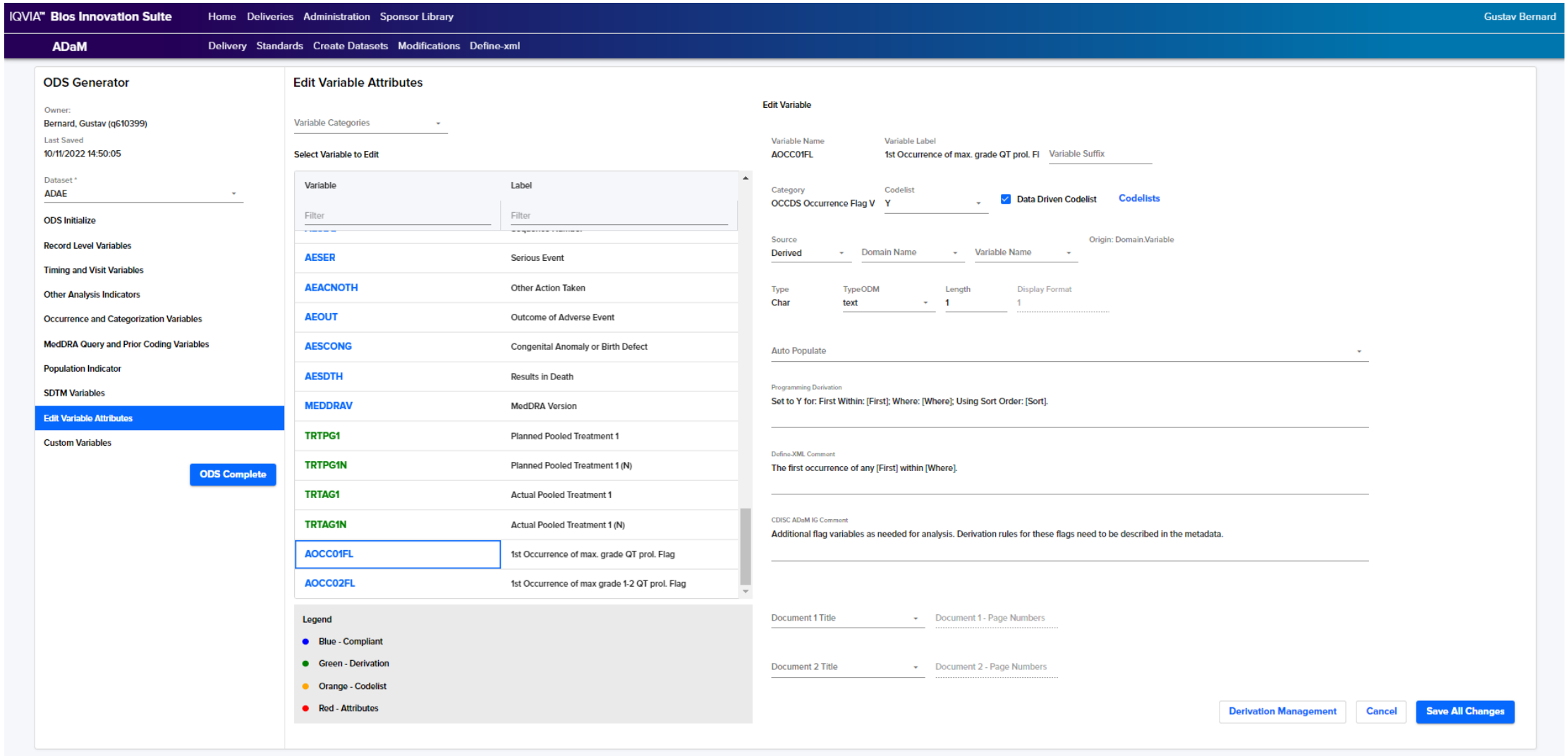
Maintenance (Rework)

- **Update** Derivation in Specification
- **Rerun** Production Program
- **Rerun** QC Program

Question Based Approach to include Occurrence Variable



Variable populated with Text Using Keywords



Derivation = Variable Calculation

Programming Derivation

Set to Y for: First Within: [First]; Where: [Where]; Using Sort Order: [Sort].

Define-XML Comment

The first occurrence of any [First] within [Where]

Programming Derivation

Set to Y for: First Within: AEBODSYS; Where: TRTEMFL="Y"; Using Sort Order: USUBJID AEBODSYS.

Define-XML Comment

The first occurrence of any AEBODSYS within TRTEMFL="Y"

1

2

3

Utilize Spec Metadata
Call %adam_initialize

Call %adam_flags

Identify all Flag variables

Identify if automated syntax present

Extract Metadata: variable, first, where, sort

Calculate Variable

Library Usage Numbers

- Number of **Studies** Processed 264
- Total Amount of ADaM **Datasets** 3692
- Average Amount of **Datasets Per Study** **13**
- Average Amount of **Variables Per Datasets** 83

Class Level Analysis

Trend \ Usage Analysis

CLASS	Nr	Ave Variables	ADSL Variable Ave	Custom Variables Ave	SDTM Variables Ave	ADaM
Subject Level Analysis Dataset	265	91		39.41	7.83	43.76
Occurrence Data Structure	1099	87	43.31	5.43	22.31	15.95
Basic Data Structure	2268	81	41.68	3.42	13.05	22.85
Basic Data Structure - TTE	48	57	41.77	1.54	0.17	13.52

Dataset Level Analysis

Dataset Name	Nr of Studies with Dataset	Source = Assigned	Source = Derived	Source= Predecessor
ADSL	265	8	57	25
ADAE	257	5	31	80
ADCM	237	2	14	72
ADVS	237	8	17	55
ADLB	233	12	27	63
ADMH	228	1	5	60
ADEG	197	9	19	56

Individual Dataset Level Analysis

Trend \ Usage Analysis

Dataset Name	Amount of Studies with Domain	% of Studies with Domain	Overall Ave Variables Per Domain	Source_Assigned	Source_Derived	Source_Predecessor	
ADLB	233	88	102	12	27	63	

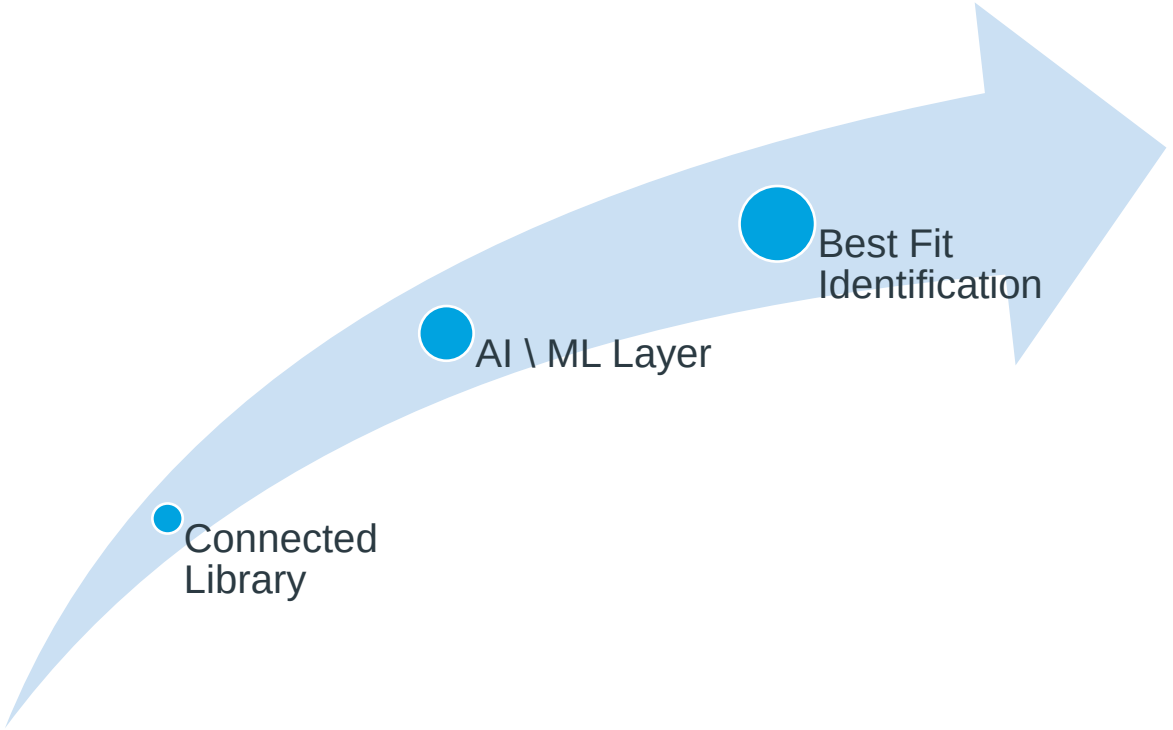
ADSL Variable	Custom Variables	SDTM Variables	Timing and Visit Variables	Parameter Variables	Baseline and Change from Baseline Variables	Record-level Variables
40.61	5.52	21.51	6.61	7.84	4.17	3.52

Analysis Criteria and Indicators	Dataset Initialize	Other Analysis Indicators	Custom Specified - SDTM Input	Custom Specified - ADaM Availible	Toxicity and Range Variables
5.61	.81	.04	.79	1.85	5.30

Study Library

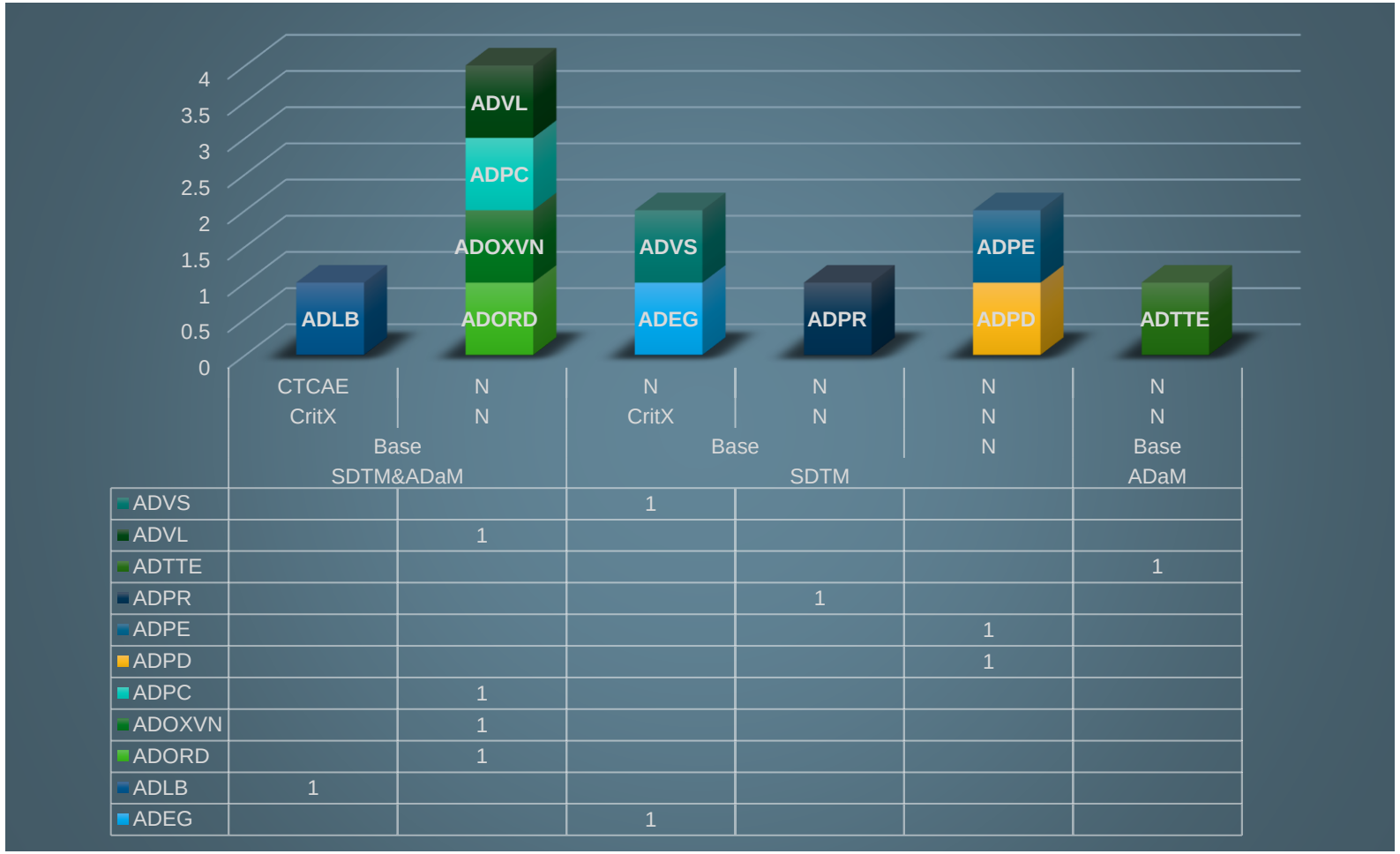
- Each Completed Study becomes a Library Project that can be Reused by any other Study.
- Client Specific Libraries
- Safety Libraries
- Therapeutic Area Libraries

- Using Libraries to startup code
- After having enough connected metadata (500+ studies) can start using AI \ ML to help identify best fit to startup new studies



Building the COVID19 Library

SDTM VLM Metadata + TLF VLM Metadata → Classify BDS ADaM Datasets



Inputs for Domain

a) Parameters Assigned from SDTM Only

b) Parameters Assigned from SDTM and New Derived Tests

c) New Derived Parameters Only

Baseline Variables Needed

a) Yes

b) No

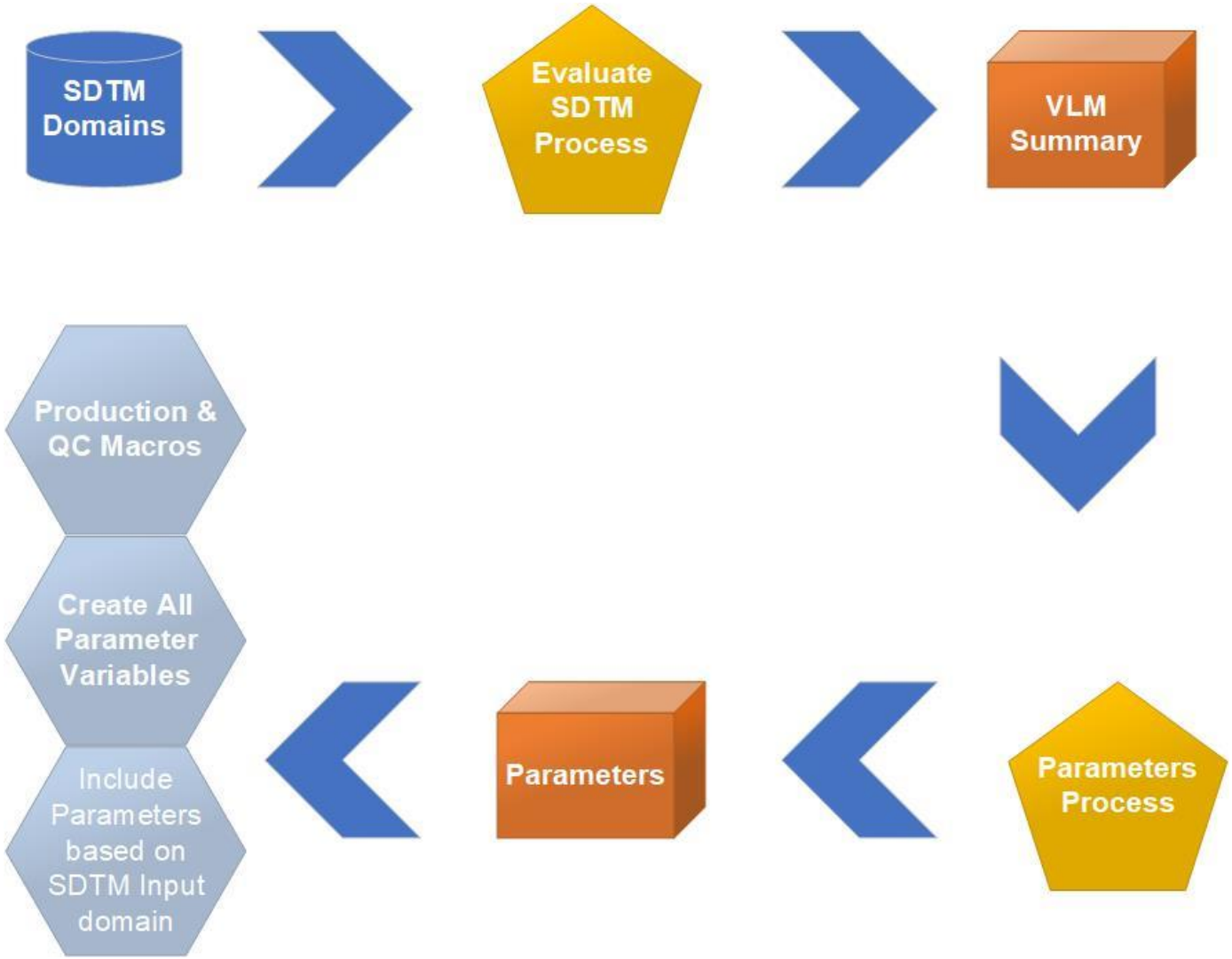
Criteria or CTCAE Variables Needed

a) Yes

b) No

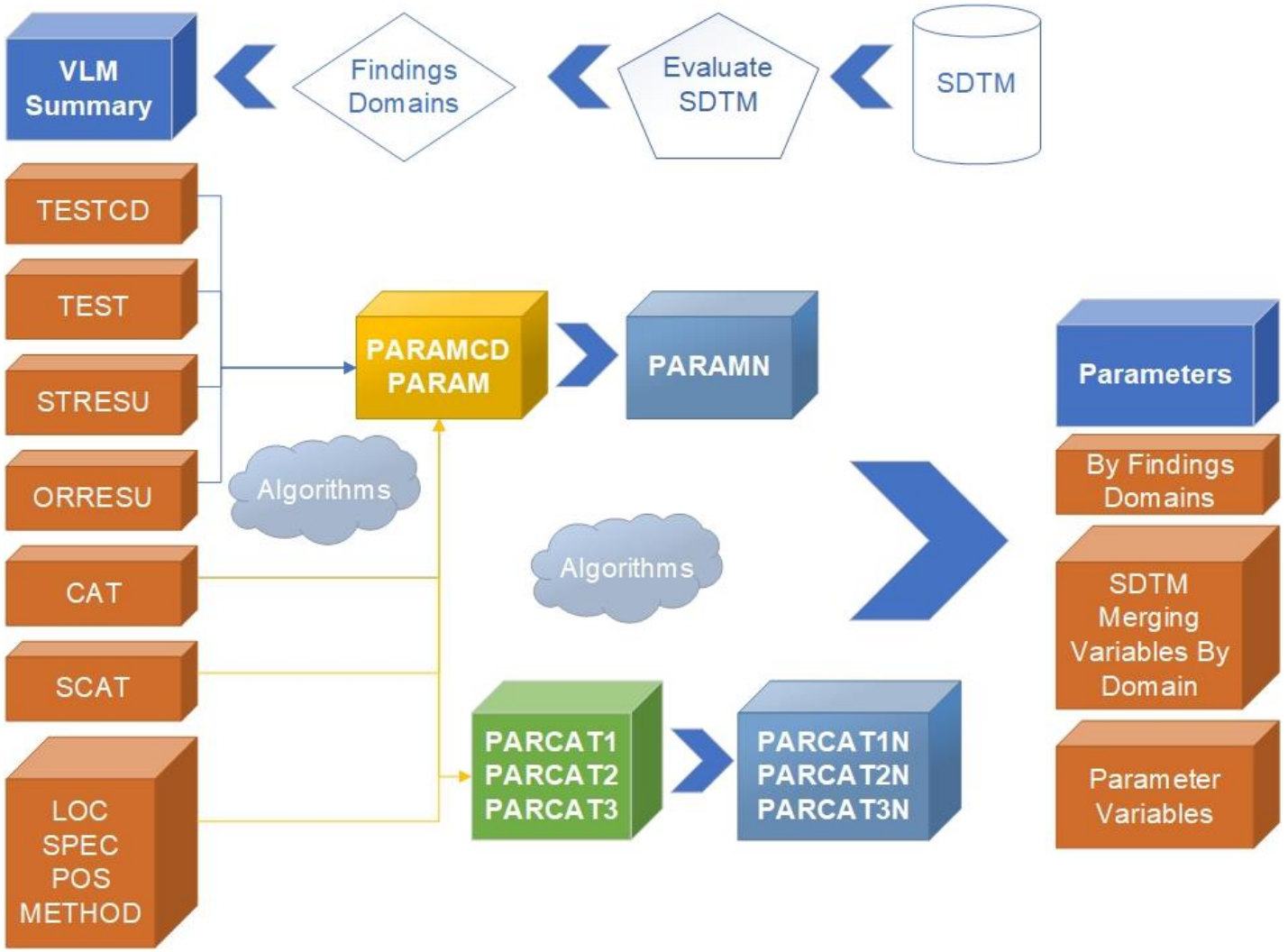
Automation Processes

Parameter Process



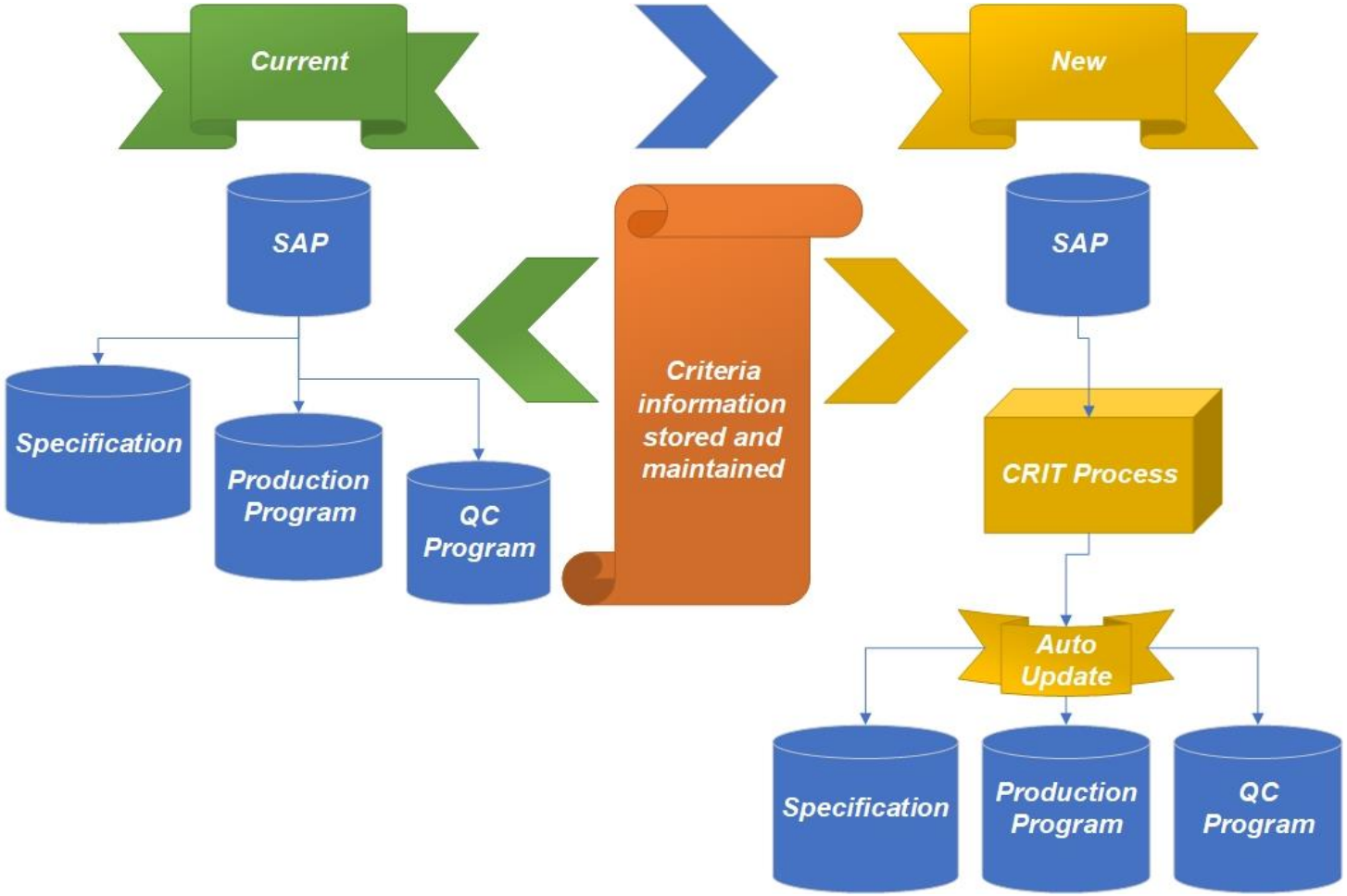
Process to Create Parameter Variables (PARAMCD, PARAMN, PARAM, PARCAT1, PARCAT1N, PARCAT2, PARCAT2N, PARCAT3 and PARCAT3N) using a Data Driven approach along with Algorithms to ensure Consistency.

Parameter Process



Automation Processes

CRIT Process

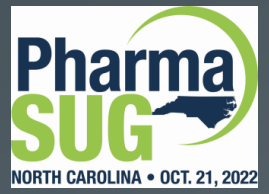


Process to Decrease the amount of Files that Store and Maintain Criteria information.

ADaM Automation Roadblocks

Common Considerations

- CDISC Define-XML Standard not very set
- CDSIC ADaM Standards multiple ways to do the same thing
- FDA requesting sas \ r code
- When Automating SAS code
  - Need to still be able to submit to FDA
  - Spec needs to still be in adequate for
    - › Sponsor Review
    - › Create Define-xml
    - › Programming Needs
    - › Code Automation



Q&A

- Gustav.Bernard@iqvia.com