

DEFINING STANDARDS: CHALLENGES AND SOLUTIONS IN THE DEVELOPMENT OF DEFINE-XML FOR SDTM AND ADAM

SOFIA VALE, OCS LIFE SCIENCES, THE NETHERLANDS

CHALLENGES

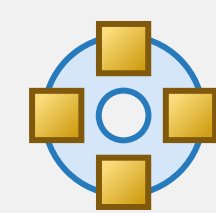
Producing Define-XML requires a lot of manual input which is **time-consuming** and **prone to errors**.

There are solutions in the market but they are often **expensive** and **inflexible**.

OUR SOLUTION

BUILD YOUR OWN!

DATA SOURCES



SAS Studio Task



Trial data or shells



Annotated CRF



CDISC Library API

ACCESSING THE CDISC LIBRARY API

```
/******  
Macro to access the CDISC Library API.  
******/  
%macro cdisc_library(url=, outlib=);  
  
    filename cdisclib temp;  
  
    proc http url = "&url."  
        auth_basic  
        webusername = "ocuser"  
        webpassword = "youwish"  
        out = cdisclib;  
    headers "Accept" = "application/json";  
    run;  
  
    libname &outlib. json fileref = cdisclib;  
  
%mend cdisc_library;  
  
/******  
Get the CDISC Library components and links to new hierarchies.  
******/  
%cdisc_library(url = http://library.cdisc.org/api/mdr/products  
    ,outlib = cdisc);
```

AND THE FUTURE?

One global task

There are currently different tasks for each standard. **One single task will be created** where you can choose the standard applicable.

Vendor independence

This solution relies on Pinnacle 21 Community to produce Define-XML. A future version will have its own engine to produce the XML file.

Analysis Results Metadata

ARM is currently not supported. A future version will create ARM in a similar fashion.

Define-XML v2.1

At the moment, the program only creates Define-XML v2.0. We are now working on support for v2.1 as well.

Custom methods and comments

Only a few fixed methods and comments are built in. A future version will support customised methods and comments.

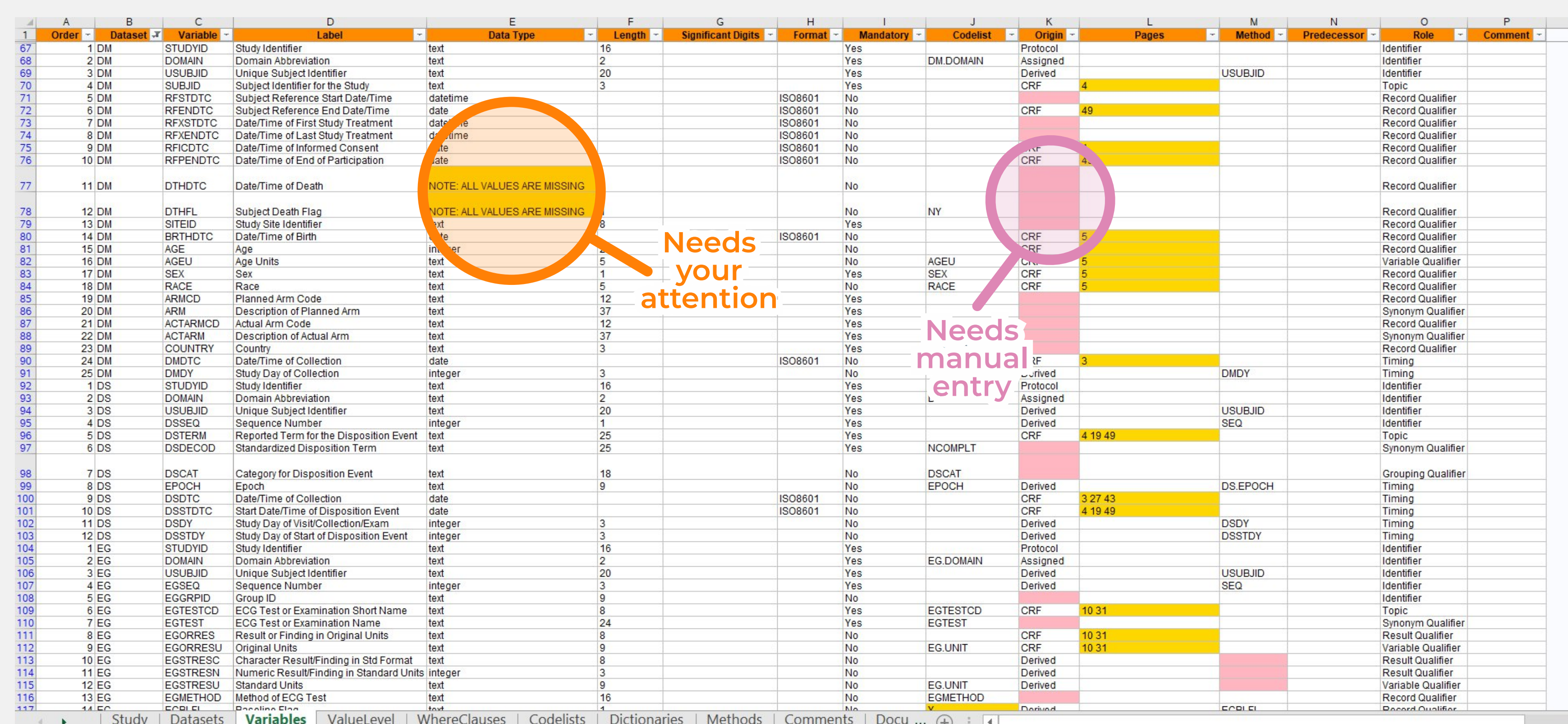
CURRENT STATUS — WHAT CAN WE AUTOMATE?

LEVEL	FULL AUTOMATION	PARTIAL AUTOMATION
Study	- Study ID - Study Description - Define Version - Standard Name and Version	
Datasets	- Dataset Name - Domain Description - Repeating - IsReferenceData - Class	Structure
Variables	- Variable Name - Length - Significant Digits - DisplayFormat - Label - Order - Key Sequence	- Data Type - Origin Type - Codelist ID - Method ID - Comment ID - Mandatory and Role
Other	Documents Metadata	- Value Level Metadata (similarly to variable level) - Codelists and Dictionaries identified in Standards - User-defined Codelists

OUTPUT

SAS programs produce an Excel spreadsheet with a structure identical to output from Pinnacle 21 Community. No need to code in XML. Highly structured and very user-friendly.

All text is automatically generated by the program. Text that is not highlighted does not need review.



Order	Dataset	Variable	Label	Data Type	Length	Significant Digits	Format	Mandatory	Codelist	Origin	Pages	Method	Predecessor	Role	Comment
67	1 DM	STUDYID	Study Identifier	text	16			Yes		Protocol				Identifier	
68	2 DM	DOMAIN	Domain Abbreviation	text	2			Yes		Assigned				Identifier	
69	3 DM	USUBJID	Unique Subject Identifier	text	20			Yes		Derived				Identifier	
70	4 DM	SUBJID	Subject Identifier for the Study	text	3			Yes		CRF				Topic	
71	5 DM	RFSTDTC	Subject Reference Start Date/Time	datetime			ISO8601	No		CRF				Record Qualifier	
72	6 DM	RFENDTC	Subject Reference End Date/Time	date			ISO8601	No		CRF				Record Qualifier	
73	7 DM	RFSTDTCT	Date/Time of First Study Treatment	datetime			ISO8601	No		CRF				Record Qualifier	
74	8 DM	RFENDTCT	Date/Time of Last Study Treatment	datetime			ISO8601	No		CRF				Record Qualifier	
75	9 DM	RFSDTC	Date/Time of Informed Consent	date			ISO8601	No		CRF				Record Qualifier	
76	10 DM	RFENDTC	Date/Time of End of Participation	date			ISO8601	No		CRF				Record Qualifier	
77	11 DM	DTHTC	Date/Time of Death	datetime				No						Record Qualifier	
78	12 DM	DTHTC	Subject Death Flag	text				No		NY				Record Qualifier	
79	13 DM	SITEID	Study Site Identifier	text	8			Yes		CRF				Record Qualifier	
80	14 DM	BRTHDTC	Date/Time of Birth	datetime			ISO8601	No		CRF				Record Qualifier	
81	15 DM	AGE	Age	integer	5			No		CRF				Record Qualifier	
82	16 DM	AGEU	Age Units	text	5			No		CRF				Record Qualifier	
83	17 DM	SEX	Sex	text	1			No		CRF				Record Qualifier	
84	18 DM	RACE	Race	text	5			No		CRF				Record Qualifier	
85	19 DM	ARMCD	Planned Arm Code	text	12			Yes		CRF				Record Qualifier	
86	20 DM	ARM	Description of Planned Arm	text	37			Yes		CRF				Synonym Qualifier	
87	21 DM	ACTARMCD	Actual Arm Code	text	12			Yes		CRF				Record Qualifier	
88	22 DM	ACTARM	Description of Actual Arm	text	37			Yes		CRF				Synonym Qualifier	
89	23 DM	COUNTRY	Country	text	3			Yes		CRF				Record Qualifier	
90	24 DM	DMTCT	Date/Time of Collection	datetime			ISO8601	No		CRF				Record Qualifier	
91	25 DM	DMDY	Study Day of Collection	integer	16			Yes		CRF				Record Qualifier	
92	1 DS	STUDYID	Study Identifier	text	16			Yes		CRF				Record Qualifier	
93	2 DS	DOMAIN	Domain Abbreviation	text	2			Yes		CRF				Record Qualifier	
94	3 DS	USUBJID	Unique Subject Identifier	text	20			Yes		CRF				Record Qualifier	
95	4 DS	DSSEQ	Sequence Number	integer	1			Yes		CRF				Record Qualifier	
96	5 DS	DSTERM	Reported Term for the Disposition Event	text	25			Yes		CRF				Record Qualifier	
97	6 DS	DSREASON	Standardized Disposition Term	text	25			Yes		CRF				Record Qualifier	
98	7 DS	DSREASON	Category for Disposition Event	text	18			No		CRF				Record Qualifier	
99	8 DS	EPOCH	Epoch	text	9			No		CRF				Record Qualifier	
100	9 DS	DMTCT	Date/Time of Collection	datetime			ISO8601	No		CRF				Record Qualifier	
101	10 DS	DSSTDTCT	Start Date/Time of Disposition Event	datetime			ISO8601	No		CRF				Record Qualifier	
102	11 DS	DSBY	Study Day of Visit/Collection/Exam	integer	3			No		CRF				Record Qualifier	
103	12 DS	DSBY	Study Day of Start of Disposition Event	integer	3			No		CRF				Record Qualifier	
104	1 EG	STUDYID	Study Identifier	text	16			Yes		CRF				Record Qualifier	
105	2 EG	DOMAIN	Domain Abbreviation	text	2			Yes		CRF				Record Qualifier	
106	3 EG	USUBJID	Unique Subject Identifier	text	20			Yes		CRF				Record Qualifier	
107	4 EG	EGSEQ	Sequence Number	integer	3			Yes		CRF				Record Qualifier	
108	5 EG	EGRPID	Group ID	text	9			No		CRF				Record Qualifier	
109	6 EG	EGTESTCD	ECG Test or Examination Short Name	text	8			Yes		CRF				Record Qualifier	
110	7 EG	EGTEST	ECG Test or Examination Name	text	24			Yes		CRF				Record Qualifier	
111	8 EG	EGORRES	Result or Finding in Original Units	text	8			No		CRF				Record Qualifier	
112	9 EG	EGORRES	Original Units	text	8			No		CRF				Record Qualifier	
113	10 EG	EGORRES	Character Result/Finding in Std Format	text	8			No		CRF				Record Qualifier	
114	11 EG	EGORRES	Numeric Result/Finding in Standard Units	text	3			No		CRF				Record Qualifier	
115	12 EG	EGORRES	Standard Units	text	8			No		CRF				Record Qualifier	
116	13 EG	EGORRES	Method of ECG Test	text	16			No		CRF				Record Qualifier	

Sofia Vale is a Statistical Programmer and Solution Developer at OCS Life Sciences in the Netherlands. She is the core developer of this solution and she is more than happy to answer any questions that you may have or to take your suggestions!



OCS Life Sciences

www.ocs-lifesciences.com

+31 (0)73 523 6000

sasquestions@ocs-consulting.com