

OpenCDISC

An open source initiative delivering
tools for validation of CDISC data



Niels Both, Principal Consultant

What is OpenCDISC?

- OpenCDISC is an open source community focused on building extensible frameworks and tools for the implementation and advancement of CDISC standards

OpenCDISC Validator

- Facilitates compliance with all CDISC models
- Supports various versions of the standards
- Provides users with control over the validation rules
- Supports near-CDISC, SDTM+, and sponsor defined standards
- Easy to integrate with existing systems and processes
- Based on freely-available open source technologies

OpenCDISC.org

The screenshot shows the OpenCDISC.org website homepage. At the top left is the OpenCDISC logo. To the right is a search bar with the text "Search". Below the logo is a navigation menu with buttons for HOME, ABOUT, VALIDATOR, RULES, DOWNLOAD, CONTRIBUTE, FORUM, and CONTACT. The main content area is divided into three columns. The left column is titled "VALIDATOR FOR CDISC" and lists three features: "Open-source", "Fully Configurable", and "Validates SDTM and more", each with a green checkmark icon. The middle column is titled "VERSION 1.1 RELEASED!" and contains two links: "Download Validator" with a green download icon and "Learn about Validator" with a green arrow icon. The right column is titled "VALIDATION RULES REPOSITORY" and contains the text "SDTM, ADaM, and other rules. Use and contribute to our open CDISC Validation Rules Repository." Below this is a section for a blog post titled "PHUSE IS COMING TO BOSTON" with a star icon and the name "MAX". The post text says "The OpenCDISC team is pleased to announce our participation in the first ever US based PhUSE Single Day Event (SDE) which will be held in Boston on September 28th, 2010 at the New England Research and Development Center. If you are not familiar with PhUSE, it's a non-profit organization dedicated to providing a forum for pharmaceutical programmers to share ideas, knowledge, and new technologies relevant to the future direction of our industry." Below the post text are links for "Delete post", "Add new comment", and "Read more". To the left of the post is a sidebar with a "SUBSCRIBE" section containing links for "Email", "Twitter", and "RSS Feed". Below the post is another section titled "OPENCDISC VALIDATOR 1.1 RELEASED" with the text "The OpenCDISC team is pleased to announce the availability of OpenCDISC Validator 1.1. Download OpenCDISC Validator 1.1 here" and a sub-section "What's New in Version 1.1" with the text "The latest release is focused on performance and scalability. The validation engine has been improved and tuned to run up to".

OpenCDISC

Search

HOME ABOUT VALIDATOR RULES DOWNLOAD CONTRIBUTE FORUM CONTACT

VALIDATOR FOR CDISC

- ✓ Open-source
- ✓ Fully Configurable
- ✓ Validates SDTM and more

VERSION 1.1 RELEASED!

- Download Validator
- Learn about Validator

VALIDATION RULES REPOSITORY

SDTM, ADaM, and other rules.
Use and contribute to our open
[CDISC Validation Rules Repository.](#)

PHUSE IS COMING TO BOSTON
MAX
Posted Thu, 08/19/2010 - 14:20

The OpenCDISC team is pleased to announce our participation in the first ever US based PhUSE Single Day Event (SDE) which will be held in Boston on September 28th, 2010 at the New England Research and Development Center. If you are not familiar with PhUSE, it's a non-profit organization dedicated to providing a forum for pharmaceutical programmers to share ideas, knowledge, and new technologies relevant to the future direction of our industry.

Delete post Add new comment Read more

OPENCDISC VALIDATOR 1.1 RELEASED
Posted Wed, 06/02/2010 - 15:13

The OpenCDISC team is pleased to announce the availability of OpenCDISC Validator 1.1.
Download OpenCDISC Validator 1.1 here

What's New in Version 1.1

The latest release is focused on performance and scalability. The validation engine has been improved and tuned to run up to

» My account
» Create content
» Administer
» Log out

SUBSCRIBE

- Email
- Twitter
- RSS Feed

Validator Page

OpenCDISC Search

HOME ABOUT **VALIDATOR** RULES DOWNLOAD CONTRIBUTE FORUM CONTACT

MAX

- » My account
- » Create content
- » Administer
- » Log out

SUPPORTED CDISC STANDARDS

- SDTM 3.1.2
- SDTM 3.1.1
- Define.xml 1.0
- ADaM 1.0 (coming soon)
- SEND 3.0 (coming soon)

SUBSCRIBE

Email

VIEW **EDIT**

OpenCDISC Validator

OpenCDISC Validator is a collaborative project to develop an open source, freely available, and commercial-quality tool for ensuring clinical data compliance with CDISC standards, including SDTM, ADaM, SEND, Define.xml, and others. The Validator's key strength is its architecture that separates validation rules from application logic. The driver behind this design decision was to provide users the ability to keep up with evolving standards, sans the need to update and re-validate application logic. The validation rules are defined using XML-based validation framework with power to specify a wide range of rules and flexibility to support near-SDTM, SDTM+, and even custom datasets.

Validator's Key Features

- ▶ Metadata driven and configurable
- ▶ Validation rules and messages stored in XML
- ▶ Facilitates compliance with all CDISC models
- ▶ Supports SDTM v3.1.2 and v3.1.1
- ▶ Checks against SDTM specification, controlled terminology, and Define.xml
- ▶ Configurable to handle SDTM+ and near-SDTM datasets
- ▶ Companies can create own standard configurations and share with partners

DOCUMENTATION

- Installing OpenCDISC Validator
- Using OpenCDISC Validator
- OpenCDISC Validation Framework
- Performance and Scalability Guide

Validation Rules Database



- HOME
- ABOUT
- VALIDATOR
- RULES
- DOWNLOAD
- CONTRIBUTE
- FORUM
- CONTACT

CDISC SDTM 3.1.2 Validation Rules

The following is a listing of CDISC SDTM version 3.1.2 validation rules supported by OpenCDISC Validator. The list includes all Janus validation rules as defined in the Janus SDTM Validation Specification v.1, CDISC controlled terminology rules, and additional rules defined by the OpenCDISC community.

Domain	Category	Type	Severity	WebSDM ID				
<input type="text"/>	<Any>	<Any>	<Any>	<input type="text"/>	<input type="button" value="APPLY"/>			
Rule ID	Message	Description of Rule	Domain	Category	Type	Severity	WebSDM ID	
CT0001	Value for --ACN not found in ACN controlled terminology codelist	Variable values should be populated with terms found in 'Action Taken with Study Treatment' (C66767) CDISC controlled terminology codelist	Events	Terminology	Warning	Medium	R5108	
CT0002	Value for AESEV not found in AESEV controlled terminology codelist	Variable values should be populated with terms found in 'Severity/Intensity Scale for Adverse Events' (C66769) CDISC controlled terminology codelist	AE	Terminology	Warning	Medium	R5110	
CT0003	Value for TSVAL not found in AGESPAN controlled terminology codelist	Variable values should be populated with terms found in 'Age Span' (C66780) CDISC controlled terminology codelist	TS	Terminology	Warning	Medium		
CT0004	Value for AGEU not found in AGEU controlled terminology codelist	Variable values should be populated with terms found in 'Age Unit' (C66781) CDISC controlled terminology codelist	DM	Terminology	Error	High	R5062	
CT0005	Value for TSVAL not found in AGEU controlled terminology codelist	Variable values should be populated with terms found in 'Age Unit' (C66781) CDISC controlled terminology codelist	TS	Terminology	Warning	Medium		

Validation Rule Details

The screenshot displays the OpenCDISC website interface. At the top left is the OpenCDISC logo. To its right is a search bar with a 'Search' button. Below the logo is a navigation menu with buttons for HOME, ABOUT, VALIDATOR, RULES, DOWNLOAD, CONTRIBUTE, FORUM, and CONTACT. The main content area is divided into a left sidebar and a main panel. The sidebar contains a 'MAX' section with a star icon and links for 'My account', 'Create content', 'Administer', and 'Log out'. Below it is a 'SUBSCRIBE' section with checkboxes for 'Email', 'Twitter', and 'RSS Feed'. The main panel shows the details for validation rule 'SD0005'. It includes 'VIEW' and 'EDIT' tabs, a 'Message' field with the text 'Non-unique value for SEQ', a 'Description of Rule' field with the text 'Identifies records where non-unique values for Sequence Number variable exist within a subject.', and fields for 'Domain: All', 'Category: Consistency', 'Type: Error', and 'Severity: High'. There is a 'Delete post' link. Below this is a table titled 'APPLIES TO' with columns for 'Standard' and 'References'. The table lists 'SDTM 3.1.1' and 'SDTM 3.1.2' under the 'Standard' column. At the bottom of the main panel is a 'COMMENTS' section with a 'POST NEW COMMENT' button.

OpenCDISC

Search

HOME ABOUT VALIDATOR RULES DOWNLOAD CONTRIBUTE FORUM CONTACT

MAX

- » My account
- » Create content
- » Administer
- » Log out

SUBSCRIBE

- Email
- Twitter
- RSS Feed

VIEW EDIT

SD0005

Message: Non-unique value for SEQ

Description of Rule: Identifies records where non-unique values for Sequence Number variable exist within a subject.

Domain: All

Category: Consistency

Type: Error

Severity: High

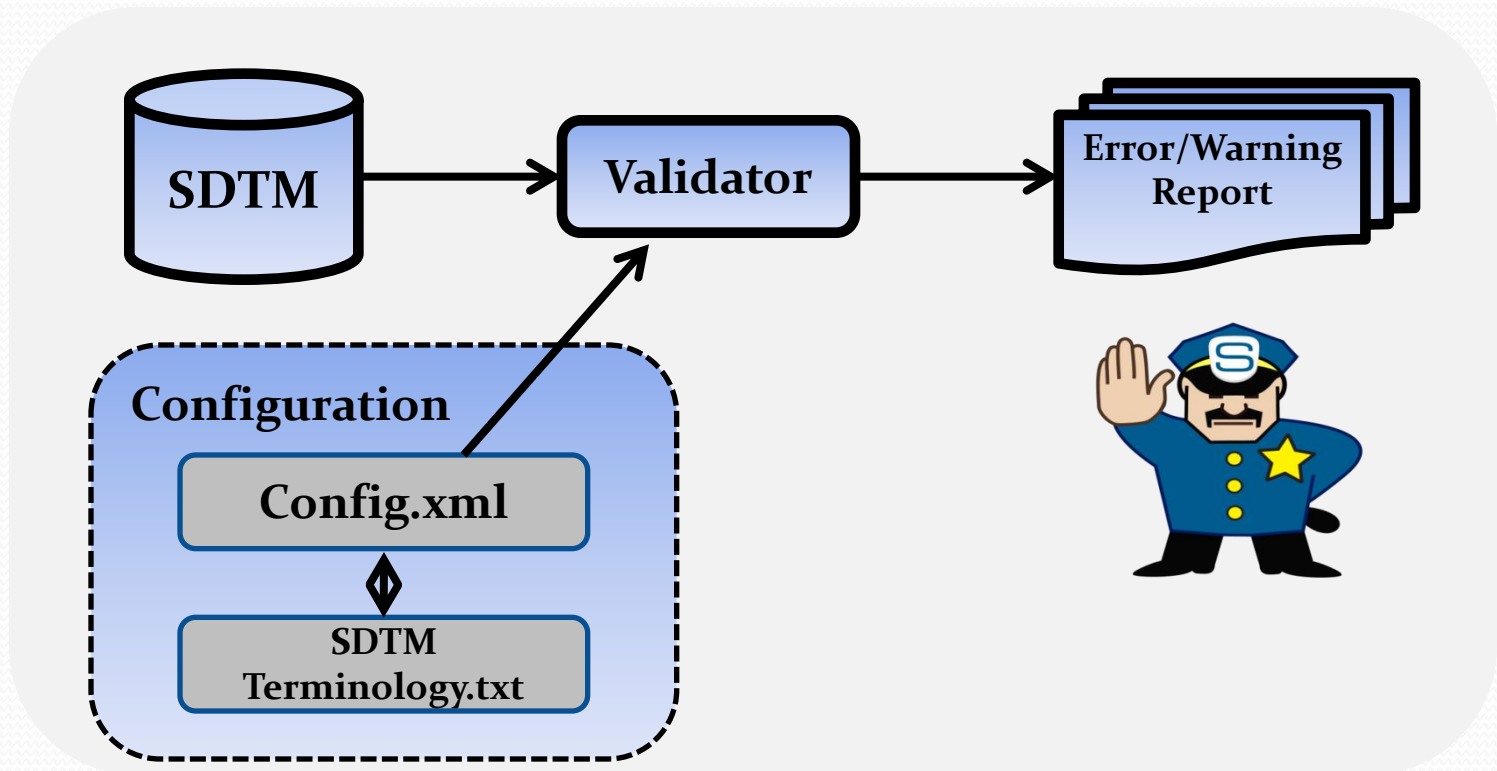
[Delete post](#)

APPLIES TO	
Standard	References
SDTM 3.1.1	
SDTM 3.1.2	

COMMENTS

POST NEW COMMENT

Validator Architecture (SDTM)



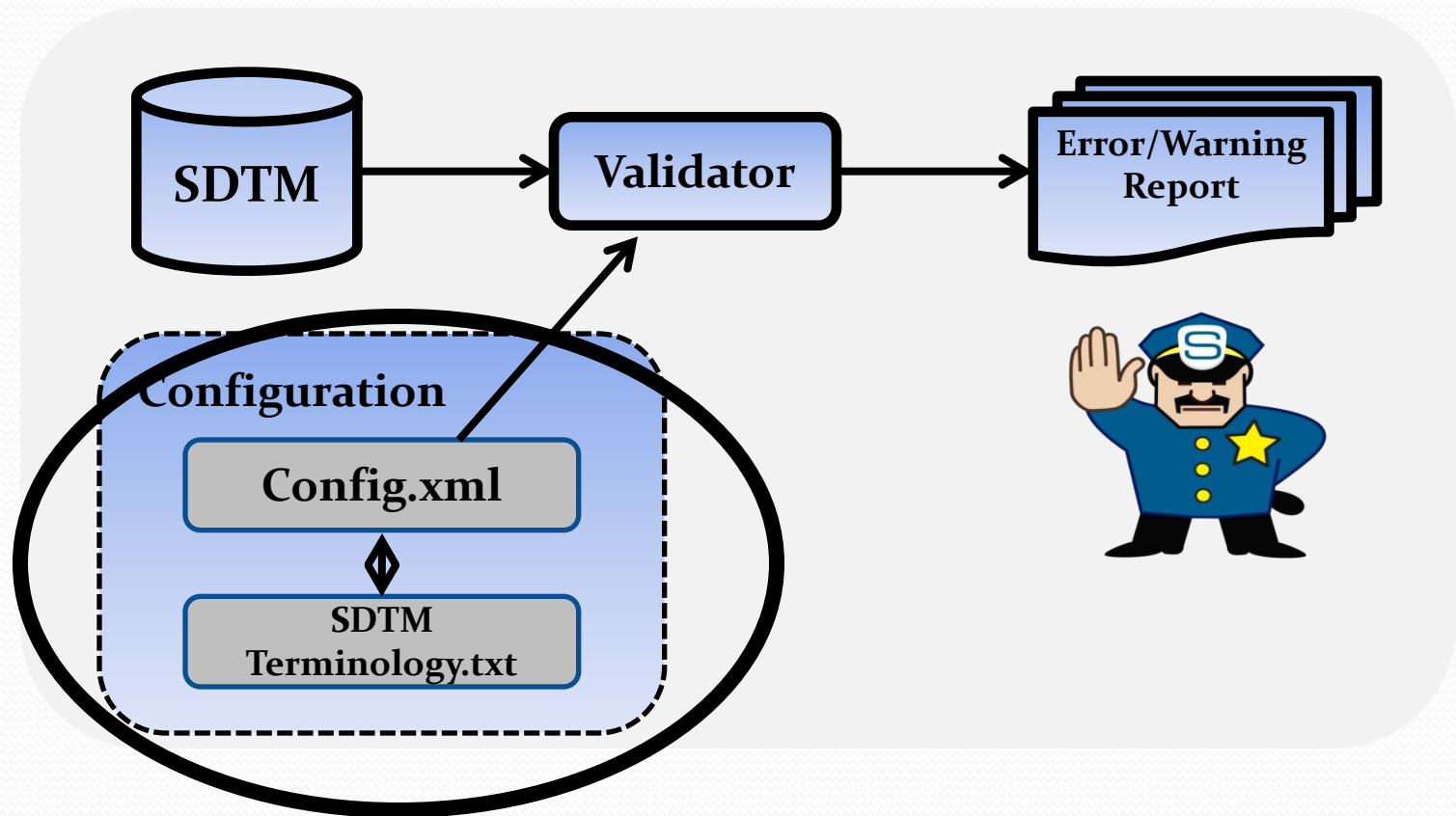
Supported Input

- SAS XPORT
- Delimited Text

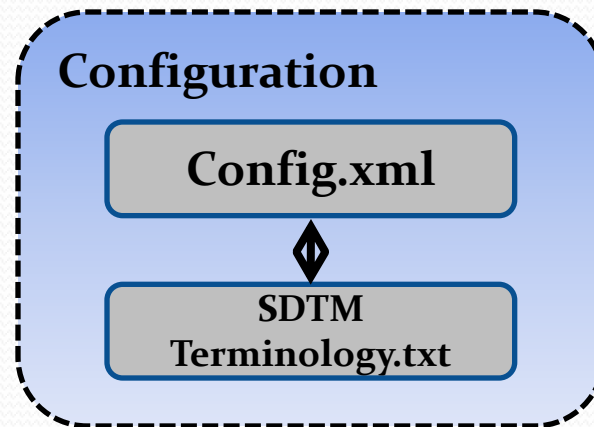
Validation Report Formats

- Excel
- CSV
- HTML
- XML

Validator Architecture (SDTM)



Configuration



Home Insert Page Layout Formulas Data Review View Add-Ins

Cut Copy Paste Format Clipboard

SDTM Terminology.txt

Wrap Text Merge & Center Alignment

General Number

A1 Code

	A	B	C	D	E	
	Code	Codelist Code	Codelist Extensible (Yes/No)	Codelist Name	CDISC Submission Value	CDISC Synonym
1	C66767		No	Action Taken with Study Treatment	ACN	Act
2	C49503	C66767		Action Taken with Study Treatment	DOSE INCREASED	
3	C49504	C66767		Action Taken with Study Treatment	DOSE NOT CHANGED	
4	C49505	C66767		Action Taken with Study Treatment	DOSE REDUCED	
5	C49501	C66767		Action Taken with Study Treatment	DRUG INTERRUPTED	
6	C49502	C66767		Action Taken with Study Treatment	DRUG WITHDRAWN	
7	C48660	C66767		Action Taken with Study Treatment	NOT APPLICABLE	NA
8	C17998	C66767		Action Taken with Study Treatment	UNKNOWN	U; U
9	C66768		No	Outcome of Event	OUT	Out
10	C48275	C66768		Outcome of Event	FATAL	Gra
11	C49494	C66768		Outcome of Event	NOT RECOVERED/NOT RESO	
12	C49498	C66768		Outcome of Event	RECOVERED/RESOLVED	
13	C49495	C66768		Outcome of Event	RECOVERED/RESOLVED WIT	
14	C49496	C66768		Outcome of Event	RECOVERING/RESOLVING	
15	C17998	C66768		Outcome of Event	UNKNOWN	U; U
16	C66780		Yes	Age Span	AGESPAN	Age
17	C27954	C66780		Age Span	ADOLESCENT (12-17 YEARS)	
18	C49685	C66780		Age Span	ADULT (18-65)	
19	C49683	C66780		Age Span	CHILDREN (2-11 YEARS)	
20	C16268	C66780		Age Span	ELDERLY (> 65)	

SDTM Terminology.txt



- Study Data Tabulation Model
- Operational Data Model
- Define.XML
- Study/Trial Design Model
- LAB
- ADaM
- Protocol
- Terminology**
- CDASH
- SEND
- CDISC SHARE
- Therapeutic Area Standards



Interchange Europe

Configuration

config-sdtm-3.1.2.xml - WordPad

File Edit View Insert Format Help



```

<val:ValidationRuleRef RuleID="SD1011" Active="Yes"/>
<val:ValidationRuleRef RuleID="SD1014" Active="Yes"/>
<val:ValidationRuleRef RuleID="SD1015" Active="Yes"/>
<!-- CDISC Controlled Terminology Rules -->
<val:ValidationRuleRef RuleID="CT0009" Active="Yes"/>
<val:ValidationRuleRef RuleID="CT0038" Active="Yes"/>
<val:ValidationRuleRef RuleID="CT0039" Active="Yes"/>
<val:ValidationRuleRef RuleID="CT0056" Active="Yes"/>
<val:ValidationRuleRef RuleID="CT0057" Active="Yes"/>
</ItemGroupDef>
<ItemGroupDef OID="SV" Name="SV" Repeating="Yes" IsReferenceData="No" Purpose="Tabulation" def:Label="Subject Visits" def:Structure=
<!-- ***** -->
<!-- Each variable is listed here for this ItemGroupDef (domain) -->
<!-- ***** -->
<ItemRef ItemOID="STUDYID" OrderNumber="1" Mandatory="Yes" Role="Identifier" val:Core="Required"/>
<ItemRef ItemOID="DOMAIN" OrderNumber="2" Mandatory="Yes" Role="Identifier" val:Core="Required"/>
<ItemRef ItemOID="USUBJID" OrderNumber="3" Mandatory="Yes" Role="Identifier" val:Core="Required"/>
<ItemRef ItemOID="SV.SVSEQ" OrderNumber="4" Mandatory="No" Role="Identifier" val:Core="Permissible"/>
<ItemRef ItemOID="SV.SVGRPID" OrderNumber="5" Mandatory="No" Role="Identifier" val:Core="Permissible"/>
<ItemRef ItemOID="SV.SVREFID" OrderNumber="6" Mandatory="No" Role="Identifier" val:Core="Permissible"/>
<ItemRef ItemOID="SV.SVSPID" OrderNumber="7" Mandatory="No" Role="Identifier" val:Core="Permissible"/>
<ItemRef ItemOID="SV.VISITNUM" OrderNumber="8" Mandatory="Yes" Role="Topic" val:Core="Required"/>
<ItemRef ItemOID="SV.VISIT" OrderNumber="9" Mandatory="No" Role="Synonym Qualifier" val:Core="Permissible"/>
<ItemRef ItemOID="SV.VISITDY" OrderNumber="10" Mandatory="No" Role="Timing" val:Core="Permissible"/>
<ItemRef ItemOID="SV.SVSTDTCT" OrderNumber="11" Mandatory="No" Role="Timing" val:Core="Expected"/>

```

Define.xml
structure

Configuration

config-sdtm-3.1.2.xml - WordPad

File Edit View Insert Format Help



```

<val:ValidationRuleRef RuleID="SD1011" Active="Yes"/>
<val:ValidationRuleRef RuleID="SD1014" Active="Yes"/>
<val:ValidationRuleRef RuleID="SD1015" Active="Yes"/>
<!-- CDISC Controlled Terminology Rules -->
<val:ValidationRuleRef RuleID="CT0009" Active="Yes"/>
<val:ValidationRuleRef RuleID="CT0038" Active="Yes"/>
<val:ValidationRuleRef RuleID="CT0039" Active="Yes"/>
<val:ValidationRuleRef RuleID="CT0056" Active="Yes"/>
<val:ValidationRuleRef RuleID="CT0057" Active="Yes"/>
</ItemGroupDef>
<ItemGroupDef OID="SV" Name="SV" Repeating="Yes" IsReferenceData="No" Purpose="Tabulation" def:Label="Subject Visits" def:Structure=
<!-- ***** -->
<!-- Each variable is listed here for this ItemGroupDef (domain) -->
<!-- ***** -->
<ItemRef ItemOID="STUDYID" OrderNumber="1" Mandatory="Yes" Role="Identifier" val:Core="Required"/>
<ItemRef ItemOID="DOMAIN" OrderNumber="2" Mandatory="Yes" Role="Identifier" val:Core="Required"/>
<ItemRef ItemOID="USUBJID" OrderNumber="3" Mandatory="Yes" Role="Identifier" val:Core="Required"/>
<ItemRef ItemOID="SV.SVSEQ" OrderNumber="4" Mandatory="No" Role="Identifier" val:Core="Permissible"/>
<ItemRef ItemOID="SV.SVGRPID" OrderNumber="5" Mandatory="No" Role="Identifier" val:Core="Permissible"/>
<ItemRef ItemOID="SV.SVREFID" OrderNumber="6" Mandatory="No" Role="Identifier" val:Core="Permissible"/>
<ItemRef ItemOID="SV.SVSPID" OrderNumber="7" Mandatory="No" Role="Identifier" val:Core="Permissible"/>
<ItemRef ItemOID="SV.VISITNUM" OrderNumber="8" Mandatory="Yes" Role="Topic" val:Core="Required"/>
<ItemRef ItemOID="SV.VISIT" OrderNumber="9" Mandatory="No" Role="Synonym Qualifier" val:Core="Permissible"/>
<ItemRef ItemOID="SV.VISITDY" OrderNumber="10" Mandatory="No" Role="Timing" val:Core="Permissible"/>
<ItemRef ItemOID="SV.SVSTDTCT" OrderNumber="11" Mandatory="No" Role="Timing" val:Core="Expected"/>

```

Activate/
Deactivate
Validation

Configuration

config-sdtm-3.1.2.xml - WordPad

File Edit View Insert Format Help



```

<!-- ***** -->
<!-- Prototype Definition section -->
<!-- ***** -->
<val:Prototype ItemGroupOID="EVENTS" KeyVariables="__TERM"/>
<val:Prototype ItemGroupOID="FINDINGS" KeyVariables="__TESTCD"/>
<val:Prototype ItemGroupOID="INTERVENTIONS" KeyVariables="__TRT"/>
<val:Prototype ItemGroupOID="SUPPQUAL" KeyVariables="QNAM"/>
<!-- ***** -->
<!-- Validation Rules section -->
<!-- ***** -->
<val:ValidationRules>
  <!-- ***** -->
  <!-- Janus Validation Rules -->
  <!-- ***** -->
  <val:Required ID="SD0002" Variable="%Variables.Core:Required%" Message="Null value in vari
  <val:Regex ID="SD0003" Variable="%Variables.Type:Date|Time%" Test="%Variable.Type.
  <val:Match ID="SD0004" Variable="DOMAIN" Terms="%Domain%" Delimiter="," Message="Inconsiste
  <val:Unique ID="SD0005" Variable="%Domain%SEQ" GroupBy="USUBJID" When="%Domain%SEQ != ' ' "
  <val:Lookup ID="SD0006" Variable="USUBJID == USUBJID" From="%Domains[EG, LB, QS, VS]%" Where="
  <val:Unique ID="SD0007" Variable="%Domain%STRESU" GroupBy="%Domain%TESTCD, %Domain%CAT" Matc
  <val:Condition ID="SD0009" Test="AESCAN == 'Y' @or AESCONG == 'Y' @or AESDISAB == 'Y' @or A
  <val:Regex ID="SD0010" Variable="VISITNUM" Test="\-?[0-9]+(\.[0-9]{1,3})?" Message="VISITNU
  
```

Validation
Rules
Definitions

Validation Rule - example

```
<val:Lookup ID="CT0002" Variable="CDISCSubmissionValue == AESEV"  
  From="FILE:TAB:%System.ConfigDirectory%/data/SDTM Terminology.txt"  
  Where="CodelistCode == 'C66769'"  
  When="AESEV != ''"  
  
  Message="Value for AESEV not found in AESEV controlled terminology  
    codelist"  
  
  Description="Variable values should be populated with terms found in  
    'Severity/Intensity Scale for Adverse Events' (C66769) CDISC  
    controlled terminology codelist"  
  
  Category="Terminology"  
  
  Type="Warning"  
  
  Severity="Medium"/>
```

How can you contribute?

- Individuals
 - Use the tools and provide feedback
 - Share your experience and provide requirements
 - Help with testing, development, or documentation
- Companies
 - Collaborate on development projects
 - Contribute code or resources

Questions/Feedback

Read paper or visit <http://www.opencdisc.org>

Presenter:

Niels Both, Principal Consultant, S-cubed

NB@s-cubed.dk

Co-author:

Max Kanevsky, OpenCDISC Founder and President

max@opencdisc.org