Tales from the journey of a CORE developer

Jozef Aerts
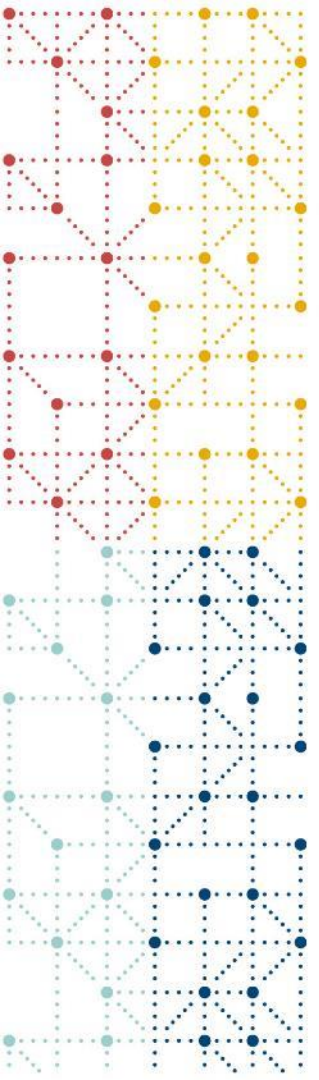**XML4Pharma**

Meet the Speaker

Jozef Aerts

Organization: XML4Pharma

Jozef is CEO of a small consultancy and software company specialized in CDISC standards. He is also a CDISC volunteer for more than 20 years and CDISC instructor.

Disclaimer and Disclosures

- *The views and opinions expressed in this presentation are those of the author(s) and do not necessarily reflect the official policy or position of CDISC.*
  (but some of them should …)

- *{Please disclose any financial relationship or conflict of interest relevant to this presentation here OR}*

- *The author(s) have no real or apparent conflicts of interest to report.*

Agenda

1. Jozef and CORE: How it started
2. CORE implementation in Software
3. Conclusions

Jozef and CORE:
How it started ...
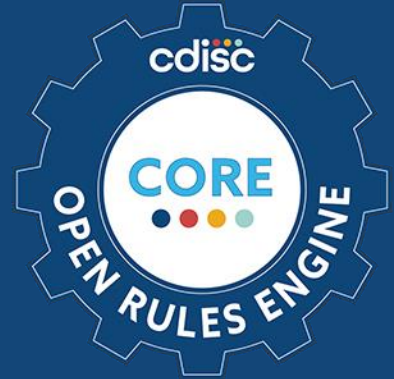
Some years ago ...

We developed the "Open Rules for CDISC Standards" (ORCS)

Executable rules were developed using XQuery and made publicly available (as Open Source)

Disadvantages:

- XQuery only works on XML

- FDA however has put the introduction of CDISC's Dataset-XML as a transport format on ice.

- Argumentation: file sizes

- Real reason: fear for change

The XML4Pharma Application Server

The "Open Rules for CDISC Standards" initiative

UPDATE! FDA SEND 3.0/3.1 validation rules 2019 (last update: 2020-07-01)
UPDATE! CDISC SDTM validation rules 2020 (v.1.1) (last update: 2020-08-23)
UPDATE! PMDA SDTM validation rules 2019 (v.2.0) (last update: 2020-08-22)
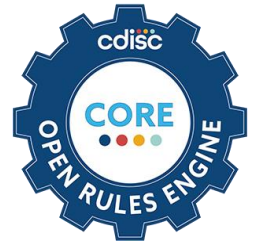*UPDATE!* FDA SDTM validation rules 2018 (last update: 2020-08-22)

These rule sets have been updated to work with define.xml 2.1.

6 years later ... CDISC CORE

How can I contribute?

- I am pretty bad in Python coding
- Many volunteers already for SDTM
- ADaM is not my thing
- Define-XML will come later

So I volunteered to implement the SENDIG (3.1) rules!

SENDIG-3.1 Rules

https://www.cdisc.org/standards/foundational/send/conformance-rules-v2-0-sendig-v3-0-and-sendig-v3-1

Development and Testing using the Rule Editor

EDIT   TEST

```
1   Check:
2     any:
3       - all:
4         - name: ARMCD
5           operator: non_empty
6         - name: ARMCD
7           operator: is_not_contained_by
8           value: $ta_armcd
9       - all:
10        - name: ARM
11          operator: non_empty
12        - name: ARM
13          operator: is_not_contained_by
14          value: $ta_arm
15  Operations:
16    - domain: TA
17      id: $ta_armcd
18      name: ARMCD
19      operator: distinct
20    - domain: TA
21      id: $ta_arm
22      name: ARM
23      operator: distinct
24  Core:
25    Id: CDISC.SENDIG.224
26    Version: '1'
27    Status: Draft
28  Description: When study does not use multi-stage arm assignments and A
29    ARMCD must be present in TA.ARMCD. This rule has been executed to ide
30    when ARMCD is not present in TA.ARMCD and therefore acknowledges that
31    may be recorded when multi-stage arm assignments are in use.
32  Outcome:
33    Message: ARMCD is not present in TA.ARMCD
34  Rule Type: Record Data
35  Sensitivity: Record
36  Authorities:
37    - Organization: CDISC
```

EDIT   TEST

✅ Validate YAML Syntax

❌ Validate YAML against Schema

✅ Convert YAML to JSON Rule

✅ Load Test Data

TEST DATASETS FILE...

Filename: unit-test-coreid-SENDIG-224-negative.xlsx

Filetype: application/vnd.openxmlformats-officedocument.spreadsheetml.sheet

Size in bytes: 13049

Last modified date: 9/27/2022, 2:20:10 PM

```
[ 2 items
  0 : { 5 items
    "filename" : "dm.xpt"
    "label" : "Demographics"
    "domain" : "DM"
    "variables" : [ 26 items
      0 : { 4 items
        "name" : "STUDYID"
        "label" : "Study Identifier"
        "type" : "Char"
        "length" : 50
      }
      1 : { 4 items
        "name" : "DOMAIN"
        "label" : "Domain Abbreviation"
        "type" : "Char"
        "length" : 50
      }
      2 : { 4 items
        "name" : "USUBJID"
        "label" : "Unique Subject Identifier"
        "type" : "Char"
        "length" : 50
      }
      3 : { 4 items
        "name" : "SUBJID"
        "label" : "Subject Identifier for the Study"
        "type" : "Char"
        "length" : 50
      }
      4 : { 4 items
        "name" : "RFSTDTC"
        "label" : "Subject Reference Start Date/Time"
        "type" : "Char"
```

✅ Results   Negatives 3   Scope Skips 1

Managing the rules in JIRA

My conclusions so far regarding Rules

- When rules were developed in the past, we did not think about computability
- Some rules express expectations (or best practices) rather than rules
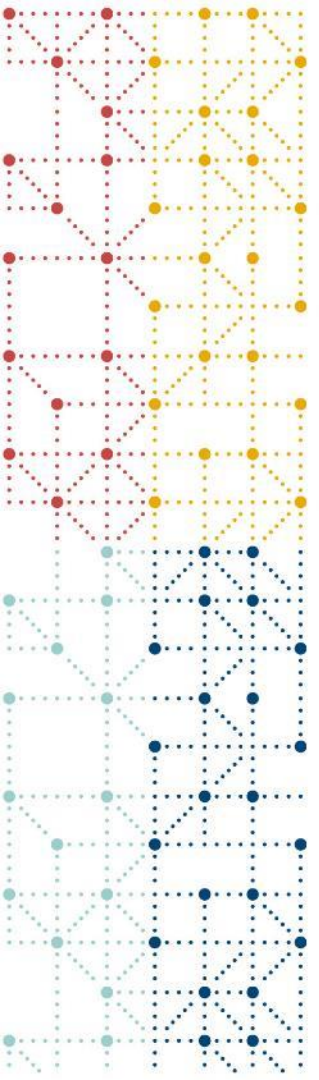
All   Comments   History   Activity

Jozef Aerts added a comment - 01/Sep/22 1:28 PM

Rule = "The values of EPOCH should provide a description of a time period that is independent of the value of ARM."

This rule is not executable as it requires semantic analysis.

- Sometimes there is "wiggle room"
- We need to get better in defining exact rules (pre-condition / post-condition)
- Implementation: avoid "*false positives*"! (under-reporting rather than over-reporting)

CORE Implementation in Software: SDTM-ETL

It's easy if you know how ...

CORE Implementation in Software
What is SDTM-ETL ?
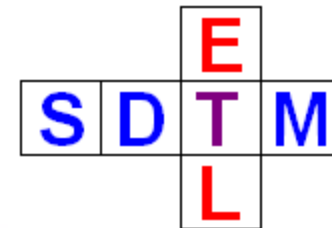

www.xml4pharma.com/SDTM-ETL

CORE Implementation in Software
SDTM-ETL, a "smart mapping tool"

- When generating SDTM/SEND datasets, one immediately wants to validate against the standard

- with the possibility to use a subset of rules only

- with the possibility to select a set of SDTM/SEND files, rather than all

- In an extremely user-friendy way ...

www.xml4pharma.com/SDTM-ETL

SDTM-ETL is now implementing CORE

Browser address bar: www.xml4pharma.com/SDTM-ETL/

NEW: Version 4.2, coming very soon!

The SDTM-ETL™ software is considered to be the lowest
clinical data in CDISC ODM format (most EDC systems do
SDTM-ETL is completely "SAS®-free", i.e. unlike other so

SDTM-ETL allows you to reuse mappings from other stud

SDTM-ETL comes with an extremely user-friendly graphic
details are to provided using intelligent wizards (no XML

Many CROs and service providers have already discovered

Latest version: SDTM-ETL 4.1

SDTM-ETL v.4.1 has full support for as well define.xml 2.

The most important documentation and a number of ma

- As of v.4.2: Validating SDTM/SEND datasets using
- SDTM-ETL 4.1 new features overview
- SDTM-ETL 4.0 new features overview

**SDTM-ETL 4.2: Validating SDTM/SEND datasets
Using CORE (CDISC Open Rules Engine)**
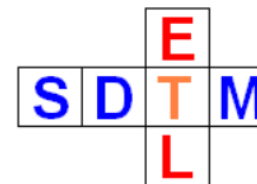
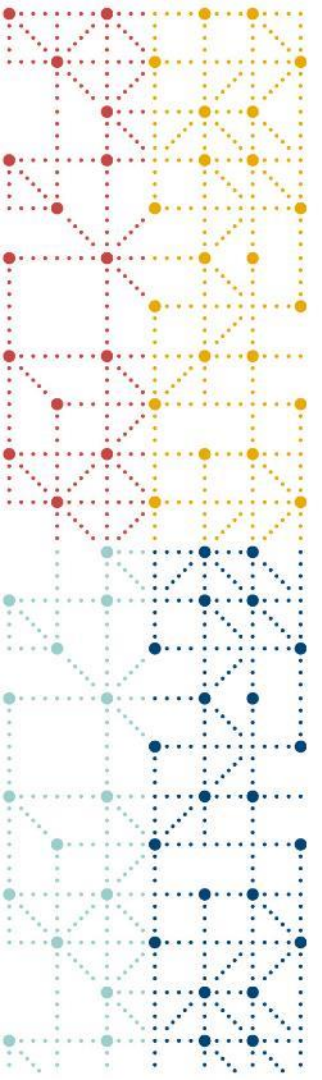Author: Jozef Aerts, XML4Pharma

Last update: **2023-04-10**

Table of Contents

Introduction

CDISC CORE (CDISC Open Rules Engine) is currently revolutionizing the submission
validation world.

CORE Implementation in SDTM-ETL: Demo movie

https://www.youtube.com/watch?v=gIgWmNN4et4

Conclusions

Conclusions

- CDISC CORE is easy to implement in software - just trigger a separate process

- As all rules and their implementations are open source and publicly available, we can finally get rid of "black box" software

- Not all rules have been implemented yet (but steadily growing)

- CORE is great for QC <u>during</u> the development of the mappings

- With CORE, every one "plays by the same rules"

- In future, CORE can be extended with company-internal rules, which makes it even more interesting
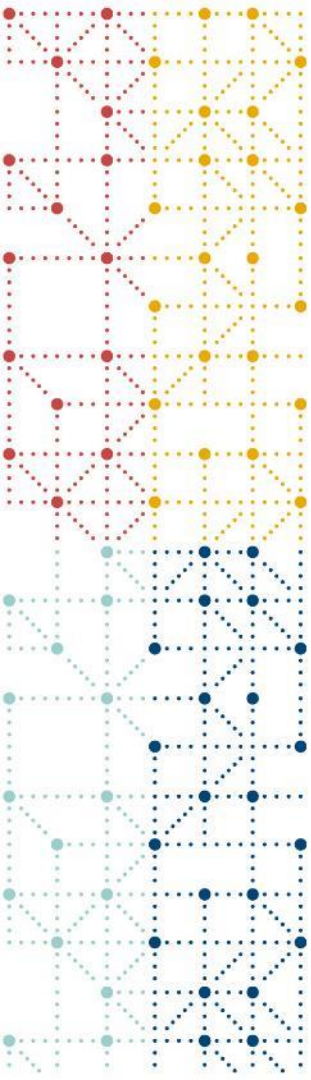
Thank You!

Contact: Jozef.Aerts@XML4Pharma.com

cdisc