2022
JAPAN
INTERCHANGE
13-14 JUNE | VIRTUAL EVENT

Upgrade/Downgrade to define v2.0 for SDTM

Hajime Shimizu, Independent (TKD+SMZ)

Meet the Speaker

Hajime Shimizu

Title: N/A

Organization: Independent (TKD+SMZ)

Hajime Shimizu works as Data Manager for more than 20 years and recently focuses on preparation of CDISC compliant datasets and metadata. In 2010, he initiated private research on CDISC standards at small group. He is known as "Mr.Akiba" in Japan CDISC community.

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.*

- *The opinions expressed in the following slides are those of the private group ("TKD+SMZ") with which the presenter participates and are not necessarily opinions of the speaker.*

- *The contents of this presentation is the result of private experiments. Please don't expect too much.*

Agenda
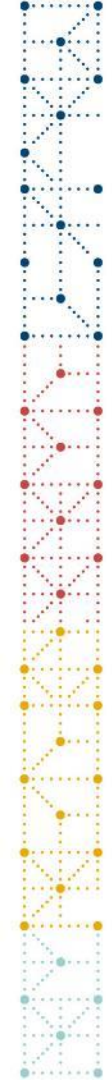
Define-XML v2.1 has come!

Rationale

- Regulatory submission

FDA	v2.1, v2.0, v1.0
PMDA	v2.0, v1.0
NMPA	v2.0?

cdisc

Experiments

Define-XML v2.1
2019-May-08

Define-XML v2.0
2014-Apr-24

Define-XML v1.0
2005-Feb-05
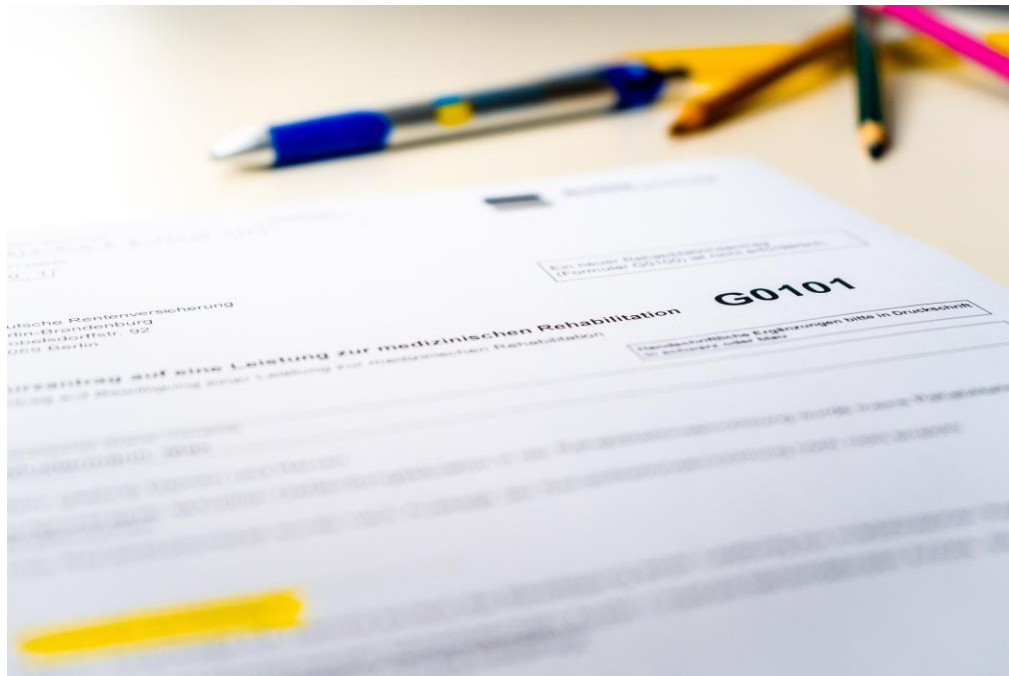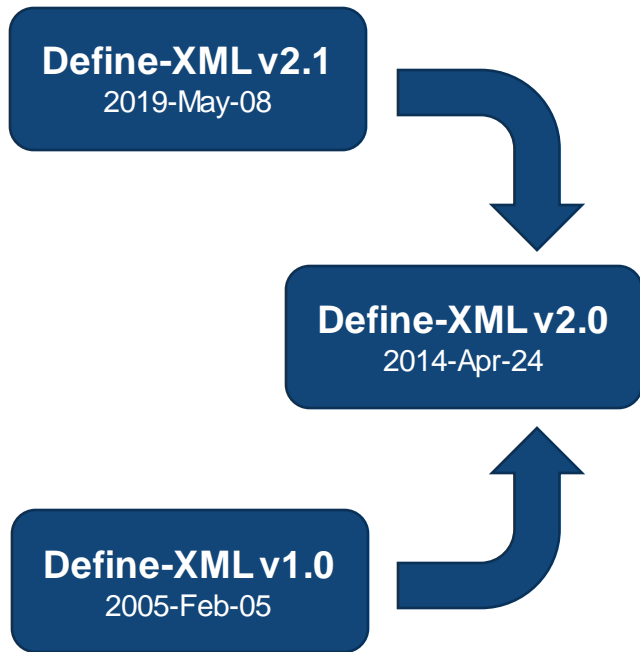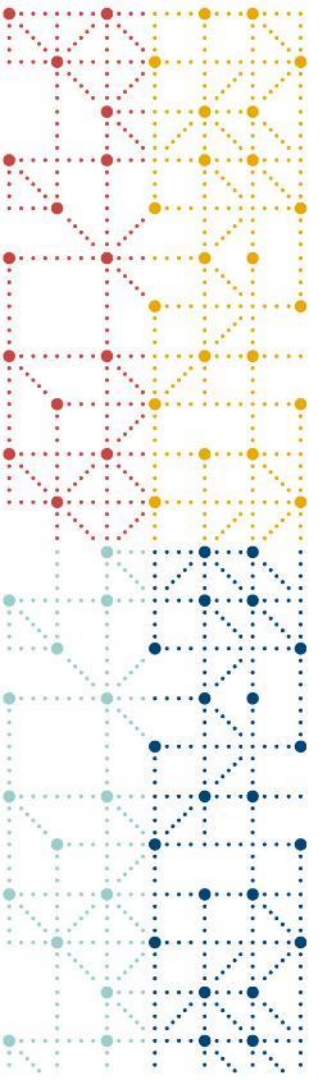
Experiment 1

Downgrade (v2.1 → v2.0)

Strategy

- Specification
  - Similar structure, contents
  - Information on v2.1 >= v2.0
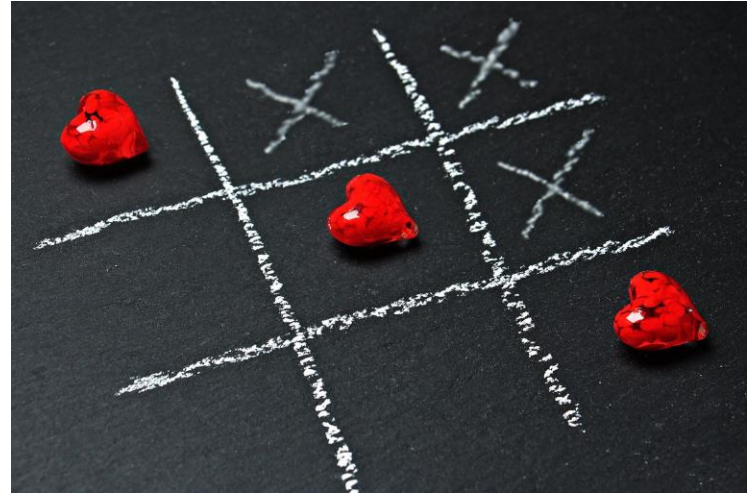
- Approach
  - Automated downgrade (Python)

- Goal
  - Pass schema validation and P21 validation
  - Minimum information loss

cdisc

Operation "Strength is the Power"

Results:

Source define	Schema	P21 (PMDA)*	Information Loss
Sample in define v2.1	Passed	Reject [2] Warning [15]	Yes
Sample in MSG v2.0	Passed	Reject [3] Warning [8]	Yes

Note:
- "SDTMIG 3.2 for PMDA" Config is applied
- Rejects/Errors/Warnings originated from source define v2.1 are found, but not listed here.

cdisc

Reject issues

- Caution to "HasNoData"
  - define v2.1 has mechanism to describe 0 record datasets
  - SDTM IG v3.4 allows to list empty datasets in define.xml
  - v3.2/3.3 prohibits to list blank datasets in define.xml

- P21 provides configs for SDTM IG v3.1.2, v3.1.3, **v3.2** and v3.3

- If empty datasets are listed in define v2.0 …
  - Triggers Reject issues
  - Hard to delete related metadata
    - Datasets, VLMs, Variables, Codelists, Terms…
    - Manual work is crazy

cdisc

2: Operation "Strength is the Power"

Results:

Source define	Schema	P21 (PMDA)*	Information Loss
Sample in define v2.1	Passed	~~Reject [2]~~ Warning [15]	Yes
Sample in MSG v2.0	Passed	~~Reject [3]~~ Warning [8]	Yes

Note:
- "SDTMIG v3.2 for PMDA" Config is applied
- Rejects/Errors/Warnings originated from source define v2.1 are found, but not listed here.

cdisc

Issues Remained

- P21 warnings are about "Leftover" fragments.
  - Comment(s)/Method(s) not referred

- Information Loss
  - Some info that can be listed in define v2.1 only is lost
    - Standards, Detailed origin, … etc.
  - If you delete orphan fragments, comment(s) will be lost
    - Need to transcribe to SDRG as appropriate

- And Issues <u>originated from source define</u> must be fixed

cdisc

Summary

- Simple program can downgrade define v2.1 automatically
  - Carefully designed define v2.1 can be downgraded in automatic

- Some corrections might be needed usually
  - Un-chained comments will be transcribed into SDRG
  - Professional review is required

- Use define v2.1 with SDTM IG v3.4
  - SDTM IG v3.4 + define v2.0
    - How shall we describe blank datasets on define? How does P21 validate?
  - SDTM IG ~v3.3 + define v2.1
    - Blank datasets can be listed on define by error

**cdisc**

What if…

- SDTM and SDTMIG Conformance Rules v2.0

Rule ID	Document	Section	Cited Guidance	Release Notes
CG0408	IG v3.2	3.2	In the event that no records are present in a dataset (e.g., a small PK study where no subjects took concomitant medications); the empty dataset should not be submitted and should not be described in the define.xml document.	
CG0408	IG v3.3	3.2	In the event that no records are present in a dataset (e.g., a small PK study where no subjects took concomitant medications); the empty dataset should not be submitted and should not be described in the Define-XML document.	
CG0408	IG v3.4	3.2	In the event that no records are present in a dataset (e.g., a small PK study where no subjects took concomitant medications), the empty dataset should not be submitted and should not be described in the Define-XML document.	No Change

cdisc

Statement in IG v3.4

- SDTM IG v3.4, Revision History

Section 3. Submitting Data in a Standard Format		
3.2	Using the CDISC Domain Models in Regulatory Submissions – Dataset Metadata	• Removed the following paragraph as there are now guidelines in the MSG v2.0 and the Define-XML v2.1 standard: "In the event that no records are present in a dataset (e.g., a small PK study where no subjects took concomitant medications), the empty dataset should not be submitted and should not be described in the Define-XML document. The annotated CRF will show the data that would have been submitted had data been received; it need not be re-annotated to indicate that no records exist."

- Maybe **false** w/z define v2.1, Maybe **true** w/z define v2.0…
- Hope to resolve this for our future (especially CORE)

cdisc
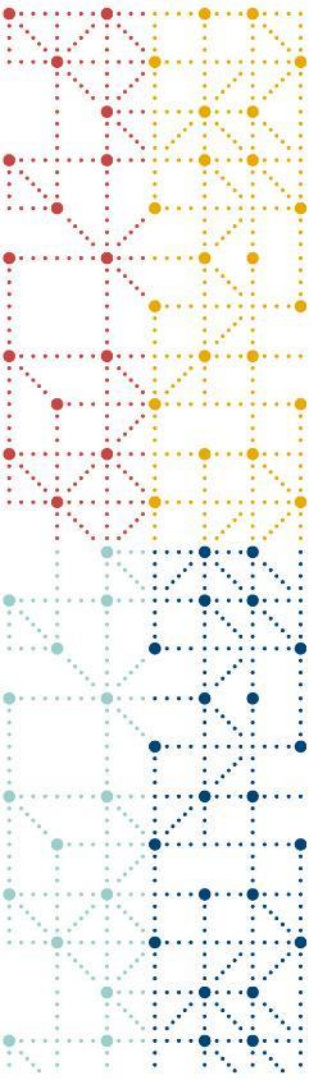
Experiment 2

Upgrade (v1.0 → v2.0)

Barriers to upgrade

- Shortly, you will see that
    - Incompatible specification
    - Information on v2.0 >> v1.0
    - Chaos in VLM implementation in v1.0
    - Poor quality of instances

- How should I handle these?

cdisc

Strategy

- "as it is" or "good define" or "better define"?

Define v2.0 as it is

- P21 supports conversion process
  - define v1.0 → spec → define v2.0
  - Easy and quick
  - Free tool

- But the output is …
  - Chaos VLM doesn't make sense
  - No new features leveraged
  - Errors are found

  - Any merits for upgrade?
  - Enough good quality to submit to PMDA?


cdisc

Good define v2.0 (Fix spec gap)

- Maximum automation is high recommended
  - Python + ts-Clinical Metadata Desktop Tools + P21
- Technical knowledge is mandatory
- …and, it requires resources

- Updated, but looks almost the same
  - Spec gap may be covered, but errors are still there
  - Is that worth doing?

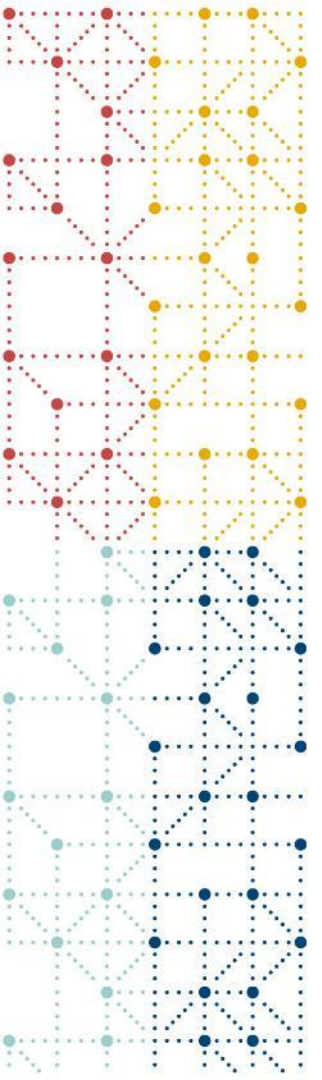**cdisc**

Better define v2.0 (Fix all gaps)

- NOT upgrade, rather full-scratch build
- Overkilling, overwork
- Not all errors can be fixed
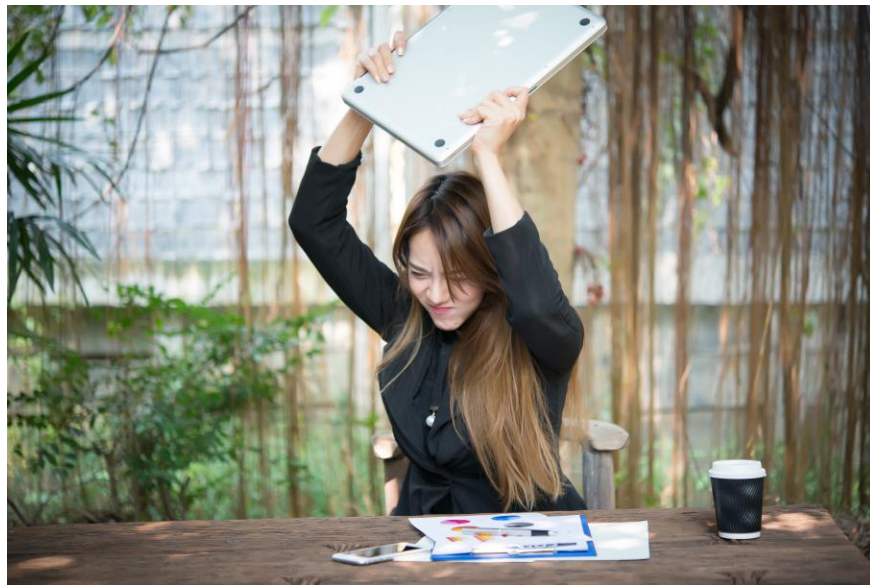
Summary

- Goal is important
  - **Philosophy** rather than **Methodology**
  - "As it is" vs. "Good one" vs. "Better one"

- Experts are needed
  - Lead
  - Technical skill
    - Create define and fix errors
    - Decipher chaos VLM
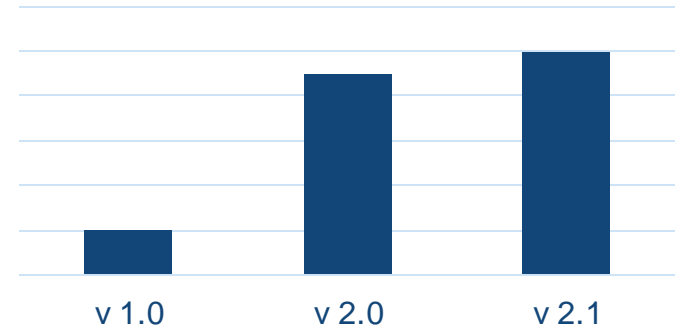   - Analysis of P21 issues
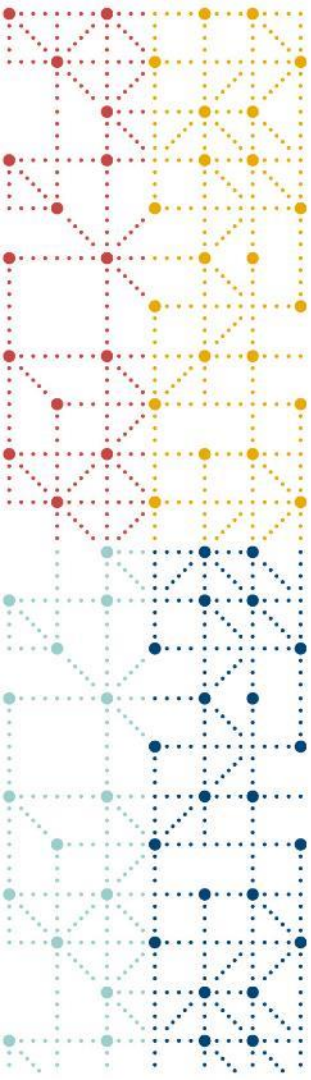

cdisc

Conclusions

Conclusion

- Downgrade is relatively easy. On the other hand, upgrade is hard.

- Specification:
  - Downgrade (v2.1 → v2.0) : Small gap
  - Upgrade (v1.0 → v2.0) : Big gap +
  - Deletion vs. Addition

- Quality of contents:
  - **Crucial** for both cases and sometimes irresolvable
  - Generally, more issues are expected in artifacts

cdisc

Conclusion

Create best define.xml for your future

Thank You!

cdisc