

# CDISC Data Standards Validation

How to insure SDTM  
compliance in a single click

**Nick De Donder**  
Project Manager  
Business & Decision Life Sciences

Tel +32 2 774 11 00  
Fax +32 2 774 11 99  
Mobile +32 476 82 20 27  
[nick.dedonder@businessdecision.com](mailto:nick.dedonder@businessdecision.com)

Sint-Lambertusstraat 141 Rue Saint-  
Lambert  
1200 Brussels

[www.businessdecision-lifesciences.com](http://www.businessdecision-lifesciences.com)

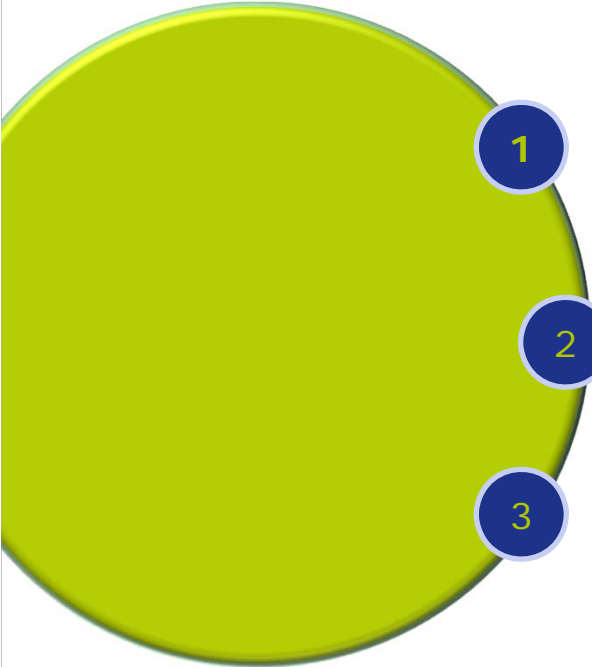


**DRIVE YOUR PERFORMANCE®**

## Topics

- 
- 1 Current Situation
  - 2 Applied CDISC Compliance
  - 3 Summary

## Topics

- 
- 1 **Current Situation**
  - 2 **Applied CDISC Compliance**
  - 3 **Summary**

## FDA Submissions (1)

- It is requested to submit clinical trial data (SDTM) & metadata (Define.xml) according to the SDTMIG V3.1.2 standards
- It is highly recommended to submit analysis data (ADaM) and metadata (Define.xml) according to the ADaMIG V1.0 standards
- Data standards allow the FDA to use standard tools

## FDA Submissions (2)

- Applies to both FDA CBER and FDA CDER (Study Data Specifications)  
<http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/FormsSubmissions/Requirements/ElectronicSubmissions/UCM163561.pdf>
- CBER has recently released new documentation for submitting data to CBER (Submission of data in CDISC format to CBER)
  - Office of Vaccines Research and Review → May 15, 2010
  - Office of Blood Research and Review → September 1, 2010
  - Office of Cellular and Gene Therapies → December 15, 2010<http://www.fda.gov/BiologicsBloodVaccines/DevelopmentApprovalProcess/ucm209137.htm>
- CDER has prepared a Data Standards Plan which will be released to the public soon  
<http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/FormsSubmissions/Requirements/ElectronicSubmissions/UCM214120.pdf>

## CDISC Data Model Conformance (SDTM & ADaM)

- Follow **complete metadata structure**
- Follow **CDISC domain models**
- Include all **Required and Expected variables**
- Use CDISC-specified:
  - standard **domain names** and prefixes
  - standard **variable names**
  - **variable labels** (SDTM only)
  - **data types** for all variables (SDTM only)
  - **controlled terminology** and **format** guidelines
- Currently, adhere to the **SAS V5 XPT limitations**
- Ensure that each dataset includes a set of **keys**
- Ensure that each record in a dataset includes a **topic variable** (SDTM only)
- Conform to all **business rules**

## Results of non-compliance

- Not following data standards will:
  - break the tools
  - create confusion and frustration
  - give the impression that the sponsor does not control the data

## Current status (1)

- Validation check history
  - WebSDM (1.5, 2.6, 3.0), SDTMIG (3.1, 3.1.1, 3.1.2), ADaMIG (1.0)
- Today, different vendors have developed a specific tool
  - WebSDM
  - Data Model Compliance Checker
  - Validator
  - Checkpoint
  - Clinical Standards Toolkit

 ORACLE | PHASE-FORWARD.

 Business & Decision  
Life Sciences

 OpenCDISC

 Octagon  
Research Solutions, Inc.

 sas

LIFE SCIENCES



## Current status (2)

- Different vendors have created different checks
- Different tools are producing different output
  - Rules and checks are not clearly defined
  - There is no official version ... yet

## Current status (3)

- ADaM team:
  - Has published ADaMIG V1.0 ( December 17, 2009)
  - Has developed a final check list (September 20, 2010)
- SDTM team did not create a check list
- CDISC has started the CAB VP
  - CDISC Advisory Board Validation Project
  - To create one official version of the SDTM checks
  - Broad industry representation
  - Representing all of the 5 major vendors and the FDA

# Current status (4)

Data Model Version		Check	Sub CheckID	Error_Message	Type	Applies to	Severity	WebSDM	WebSDM Rule Description
SDTMIG V3.1.2		1001		DATE VARIABLE does not have a ISO 8601 format.	Data	All observation classes, TD, Special Purpose	Medium	IR5002	Identifies records where the variable does not conform to the ISO 8601 standard.
SDTMIG V3.1.2		1002		The STANDARD UNIT is not consistent per LAB TEST or EXAMINATION.	Data	Findings	High	IR5006	Identifies Measurement, Test or Examination records where the standard units value (Standard Unit) is not consistent with the expected standard units value.
SDTMIG V3.1.2		1003	1	The AE is serious, but none of the serious criteria is answered 'Y'.	Data	Events	Medium	IR5008	Identifies records where Seriousness is 'Y' but none of the criteria for Seriousness (Congenital Anomaly or Birth Defect, Results in Death, Requires or Other Medically Important Study) equals 'Y'.
SDTMIG V3.1.2		1003	2	The AE is not serious, but one of the serious criteria is answered 'Y'.	Data	Events	Medium	IR5008	Identifies records where Seriousness is 'Y' but one of the criteria for Seriousness (Congenital Anomaly or Birth Defect, Results in Death, Requires or Other Medically Important Study) equals 'Y'.
SDTMIG V3.1.2		1004	1	A STANDARD RESULT is completed, but the STATUS is completed.	Data	Findings	Medium		
SDTMIG V3.1.2		1004	2	No STANDARD RESULT is completed, but STATUS is not completed.	Data	Findings	Low		
SDTMIG V3.1.2		1005		START DATE/TIME OF OBSERVATION falls after the END DATE/TIME OF OBSERVATION.	Data	All observation classes, Special Purpose	High	IR5101	Identifies records that violate the observation period constraint: Observation less than or equal to records where [Start Date/Time of Observation] is not less than or equal to [End Date/Time of Observation].
SDTMIG V3.1.2		1006		The OCCURRENCE is completed differently from 'Y' or 'N'.	Data	Events, Interventions	High	IR5106	Identifies records that violate the occurrence constraint: Occurrence is not null.
SDTMIG V3.1.2		1007		The LENGTH of the --TEST variable is more than 40 characters.	Data	Findings	High	IR5113	Identifies records that violate the test length constraint: Test or Examination Short Name is longer than 40 characters.
SDTMIG V3.1.2		1008	1	The --TESTCD starts with a number.	Data	Findings	High	IR5114	Identifies records that violate the test code constraint: Test or Examination Short Name is not null and starts with a number.
SDTMIG V3.1.2		1008	2	The --TESTCD contains special characters.	Data	Findings	High	IR5114	Identifies records that violate the test code constraint: Test or Examination Short Name is not null and contains special characters.
SDTMIG V3.1.2		1009		An ORIGINAL RESULT is completed but the STANDARD RESULT (CHARACTERISTIC) is missing.	Data	Findings	High	IR5135	Identifies records that violate the result constraint: Format is not null, limited to result units is not null. Identifies records where value is not null but the variable code list attached to a variable is null. Identifies user-defined code lists.

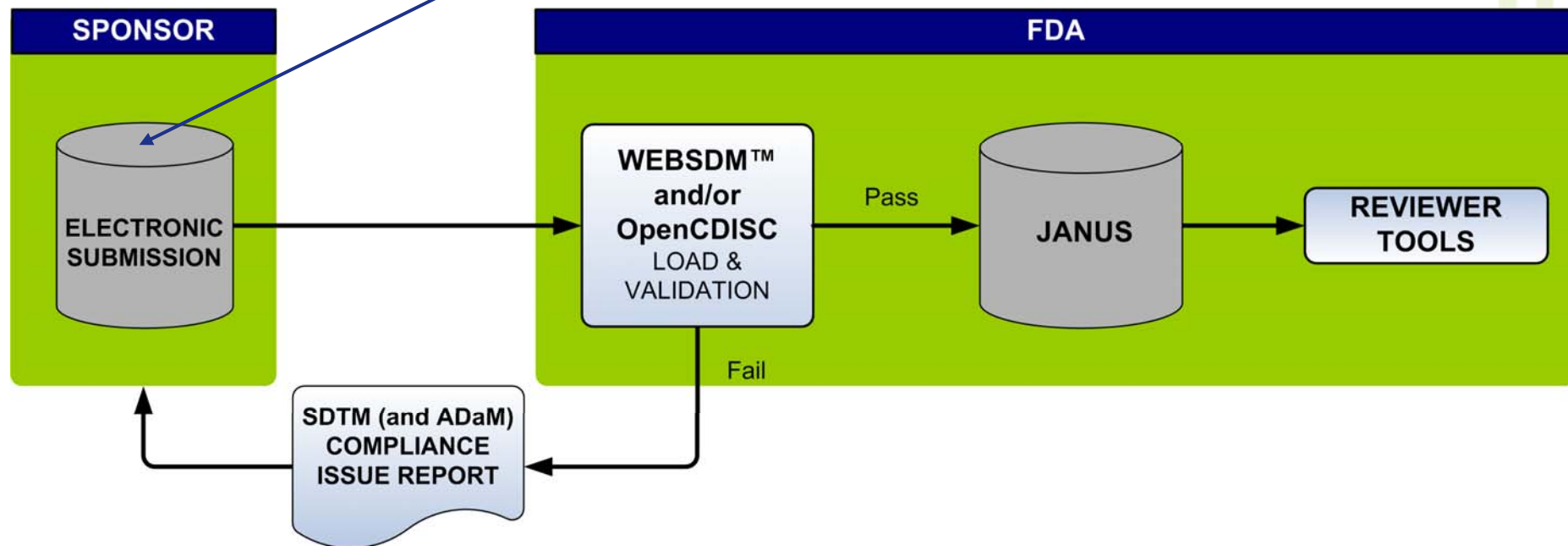
LIFE SCIENCES

- Ongoing comparison and consolidation between all checks
- Unambiguous check specification will avoid check interpretation

## FDA CRT Submission Process (1)

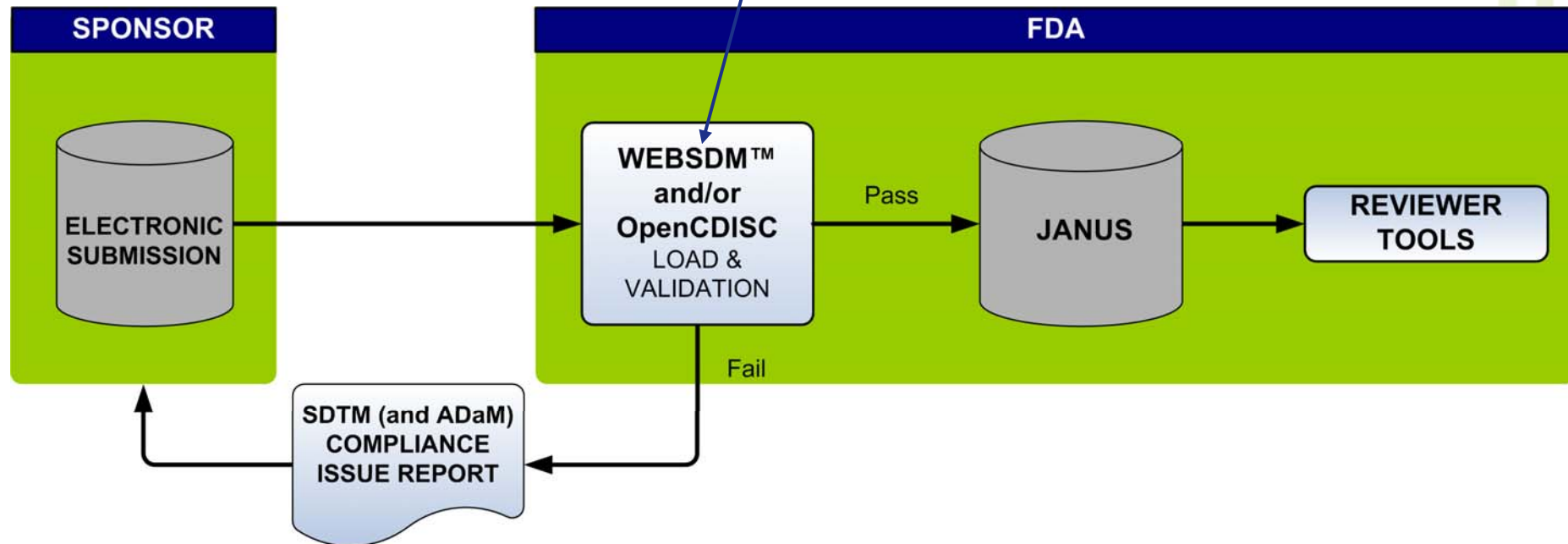
### Data Tabulation Components:

- SDTM annotated CRFs
- SDTM datasets
- ADaM datasets
- Define.xml metadata



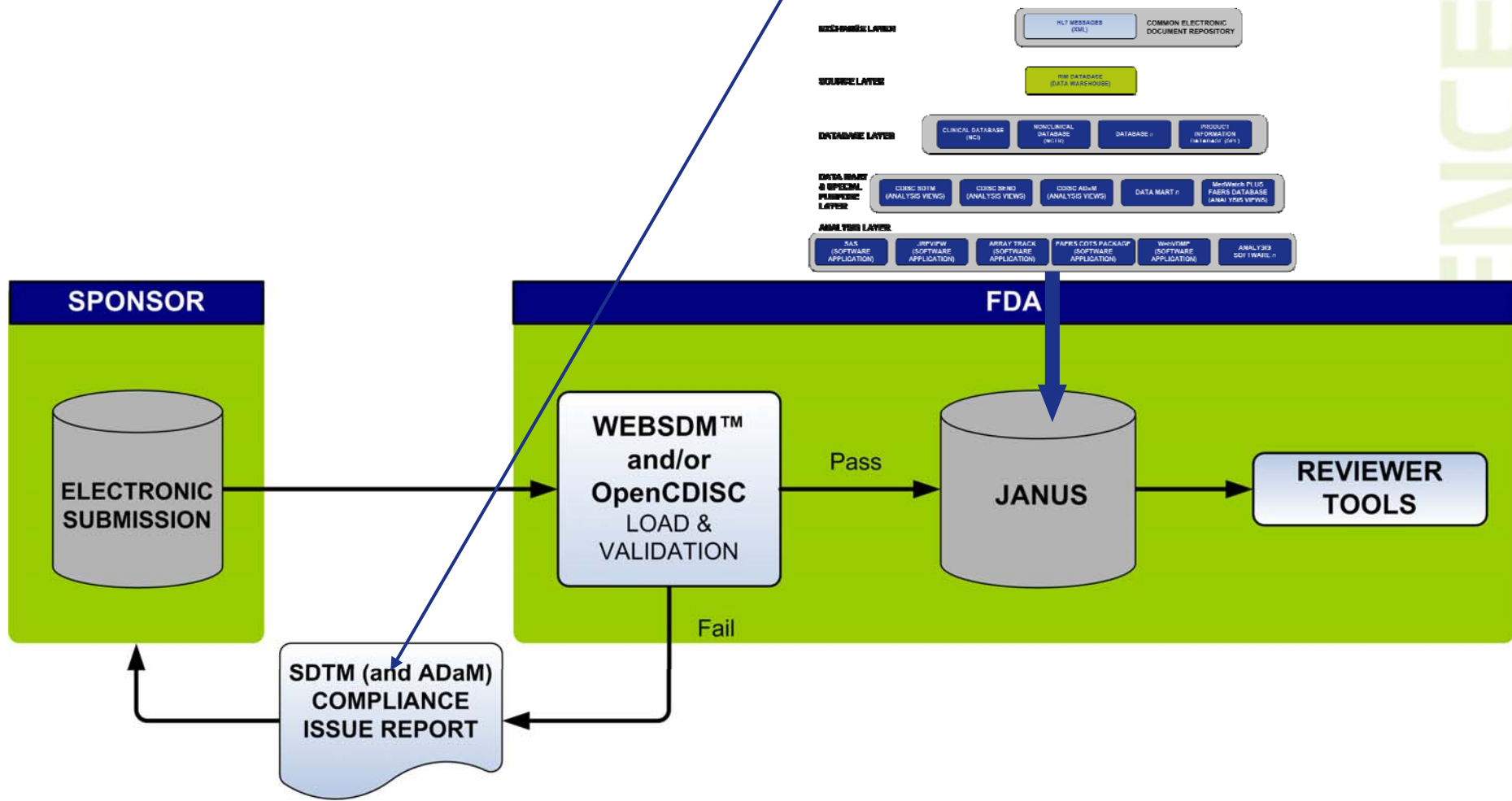
## FDA CRT Submission Process (2)

WebSDM™ and OpenCDISC tools check the CDISC SDTM data & metadata to ensure standard FDA tools can be used



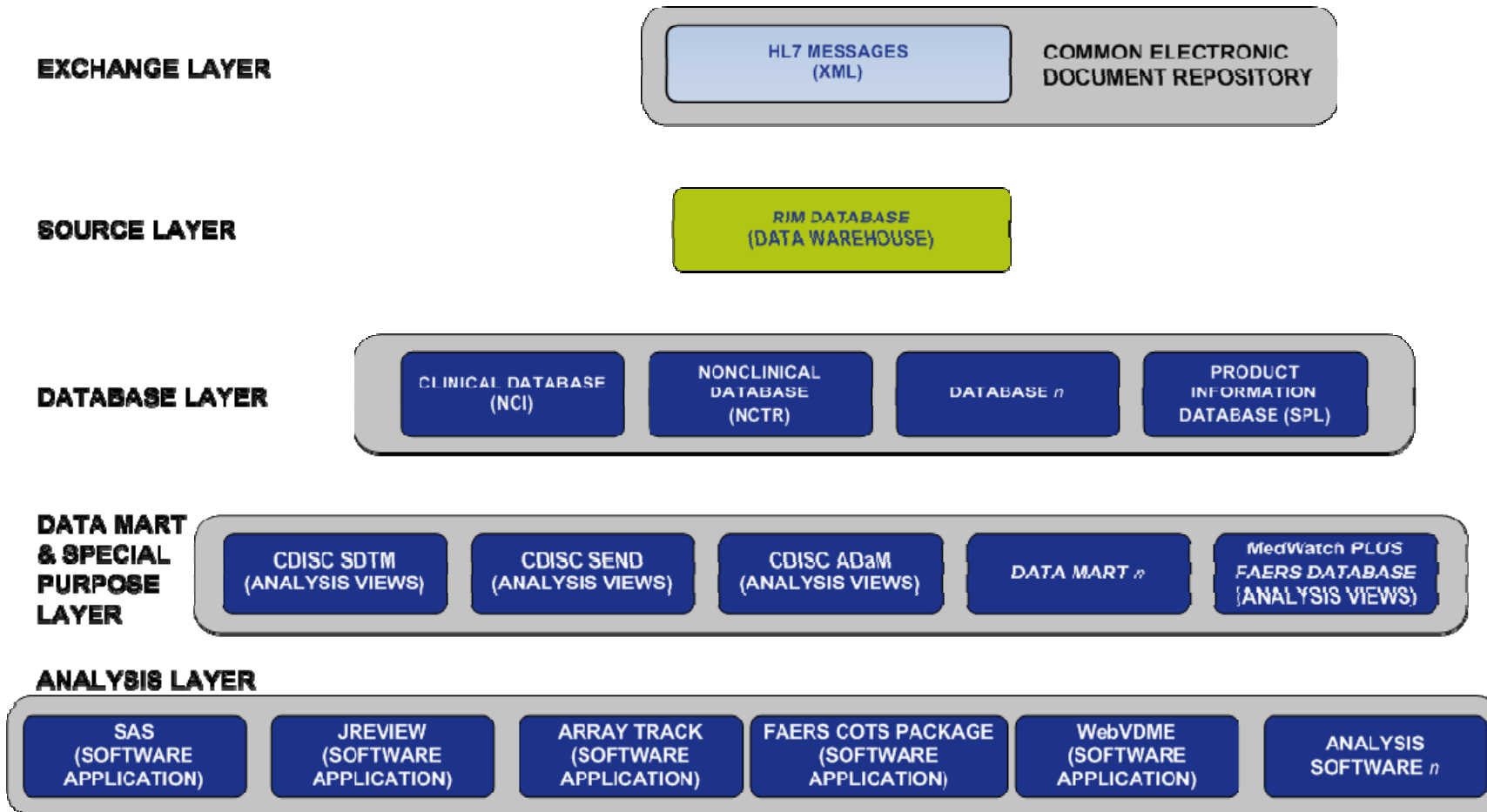
# FDA CRT Submission Process (3)

CDISC SDTM (& ADaM) compliance issues are reported to the sponsor



LIFE SCIENCES

# Janus layers



Source:

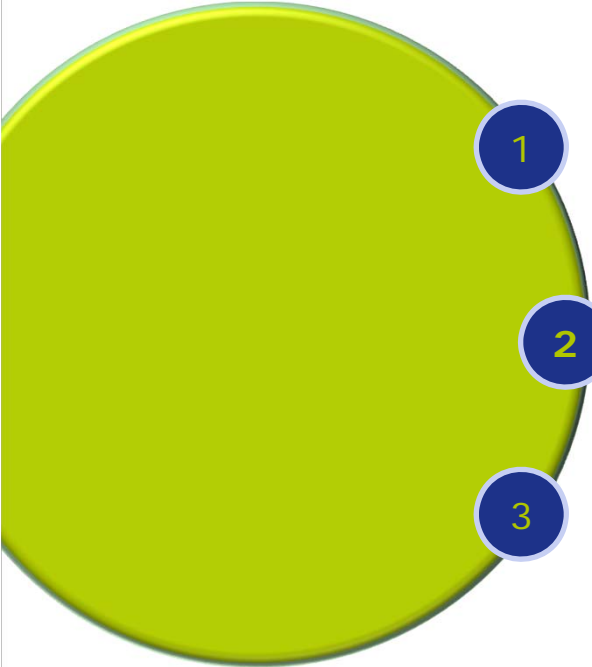
« [CDISC Standards, Present & Future](#) », Dave Iberson-Hurst, CDISC, VP Technical Strategy

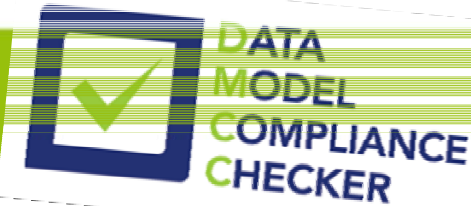
## SDTM Compliance Issues Report

- OBPS provides an SDTM issue list to the sponsor
- OBPS = Office of Business Process Support
- Loading issues can be caused by
  - Data & Metadata Inconsistency
  - Non conformity with:
    - » SDTM Model
    - » SDTM IG business rules
    - » Controlled Terminology



## Topics

- 
- 1 Current Situation
  - 2 Applied CDISC Compliance**
  - 3 Summary



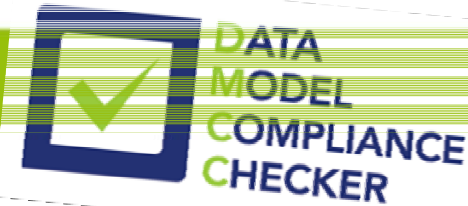
# CDISC Compliance Checks

We have created an expanded & enhanced list of checks

- 154 WebSDM™ checks
- Total check package:

	SDTMIG V3.1.1	SDTMIG V3.1.2	ADaMIG V1.0
Data checks	141	219	45
Metadata checks	68	117	51
Mapping checks	56	57	12
Project consistency checks	20	20	20

**CDISC compliance checks list is growing continuously**



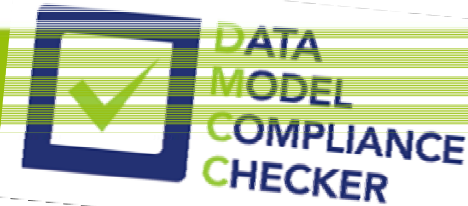
# Categorization

- CDISC Compliance Checks Categories

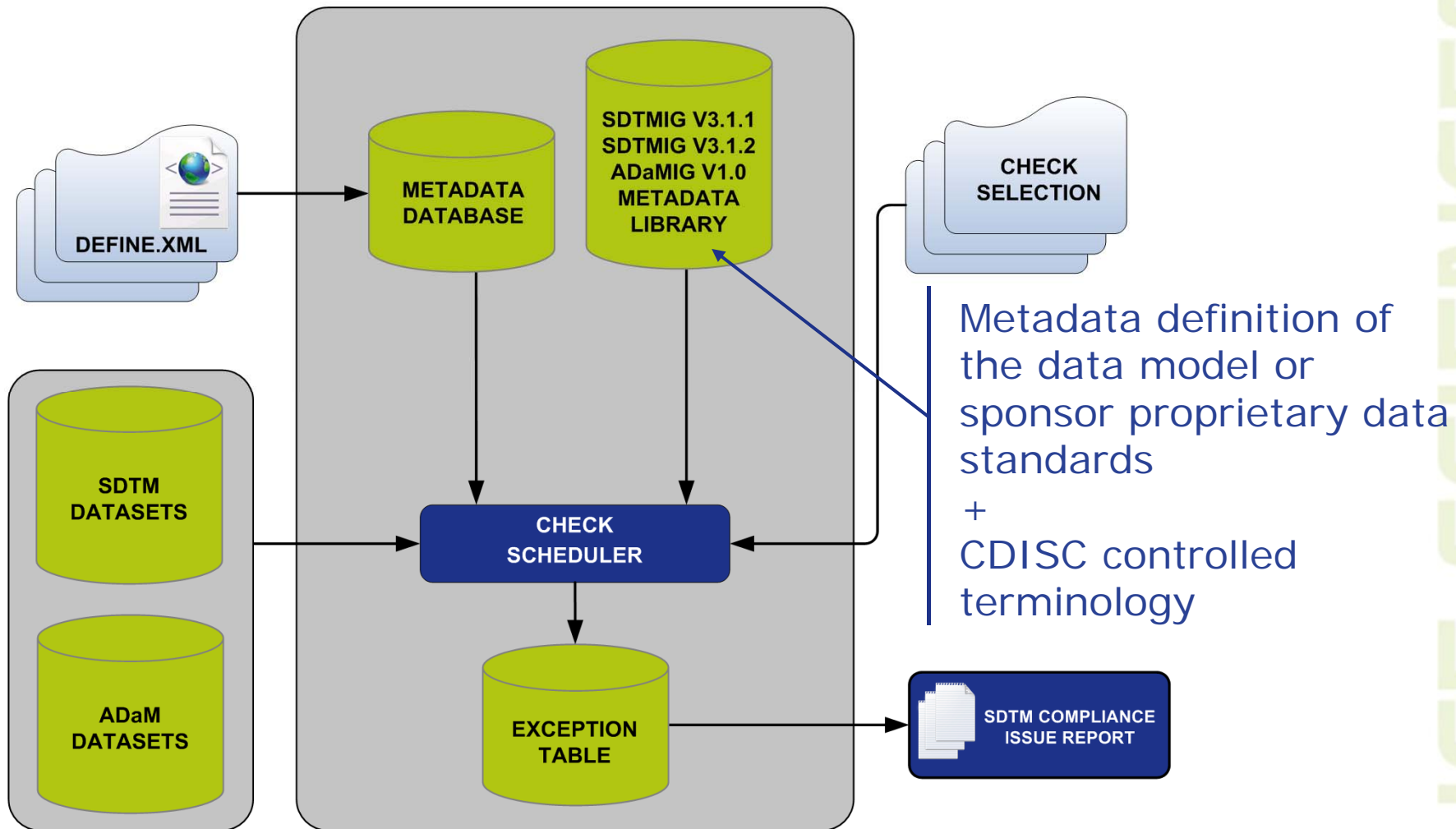
- Data
- Metadata
- Class
- Domain
- Mapping
- Data model version

CheckID	Sub_CheckID	Check_Description	Error_Message	Data_Checks	Metadata_Checks	Programming_checks	Mapping_Checks	All_Class_Checks	Events_Checks	Interventions_Checks	Finding_Checks	Trial_Design_Checks	Special_Purpose_Checks	Findings_About_Checks	Submission_Checks	Data_Model_Version
0001		Check if all the date variables (--DTC, --STDTC, --ENDTC, RFSTDTC, RFENDTC) have a ISO 8601 standard format.	DATE VARIABLE does not have a ISO 8601 format.	X	X		X	X	X	X	X	X				SDTMIG V3.1.1
0002		Check per LBTESTCD if the LBSTRESU is consistent within the trial limited to records where [LBDRRES <> NULL AND LBSTRESU <> NULL]	The STANDARD UNIT is not consistent per LAB TEST or EXAMINATION.	X							X					SDTMIG V3.1.1
0003	1	Check if AESER = 'Y' that [AESCAN='Y' or AESCONG='Y' or AESDISAB='Y' or AESDTH='Y' or AESHOSP='Y' or AESLIFE='Y' or AESMIE='Y' or AESOD='Y'] (Check only applies if serious criteria are completed).	The AE is serious, but none of the serious criteria is answered 'Y'.	X					X							SDTMIG V3.1.1
0003	2	Check if [AESCAN='Y' or AESCONG='Y' or AESDISAB='Y' or AESDTH='Y' or AESHOSP='Y' or AESLIFE='Y' or AESMIE='Y' or AESOD='Y'] that AESER = 'Y'.	The AE is not serious, but one of the serious criteria is answered 'Y'.	X					X							SDTMIG V3.1.1
0004	1	Check if --STRESC <> NULL that --STAT = NULL.	A STANDARD RESULT is completed, but the STATUS is completed.	X	X						X					SDTMIG V3.1.1
0004	2	Check if --STAT <> NULL that --STRESC = NULL.	No STANDARD RESULT is completed, but STATUS is not completed.	X	X						X					SDTMIG V3.1.1
0005		Check if [--STDTC <> NULL and --ENDTC <> NULL] that [--STDTC <= --ENDTC]] (Check the same level of precision of both dates/times).	START DATE/TIME OF OBSERVATION falls after the END DATE/TIME OF OBSERVATION.	X			X	X	X	X	X	X				SDTMIG V3.1.1
0006		Check if --OCCUR <> NULL that [--OCCUR='Y' or --OCCUR='N']	The OCCURRENCE is completed differently from 'Y' or 'N'.	X	X				X	X						SDTMIG V3.1.1
0007		Check if --TEST <> IETEST that [LENGTH(--TEST) <= 40 characters].	The LENGTH of the --TEST variable is more than 40 characters.	X	X						X					SDTMIG V3.1.1
0008	1	Check if --TESTCD, ARMCD, ETCD doesn't start with a number	The --TESTCD or ARMCD or ETCD variable starts with a number.	X							X	X	X			SDTMIG V3.1.1
0008	2	Check if --TESTCD, ARMCD, ETCD contains special characters (underscores are allowed)	The --TESTCD or ARMCD or ETCD contains special characters.	X							X	X	X			SDTMIG V3.1.1
0009		Check if [--ORRES <> NULL] that [--STRESC <> NULL]	An ORIGINAL RESULT is completed but the STANDARD RESULT (CHARACTERISTIC) is missing.	X	X						X					SDTMIG V3.1.1
0010		Check for each codelist related variable, that the value is found in the study-specific codelist attached to that variable.	The value cannot be found in the codelist attached to the variable.	X				X	X	X	X	X	X			SDTMIG V3.1.1

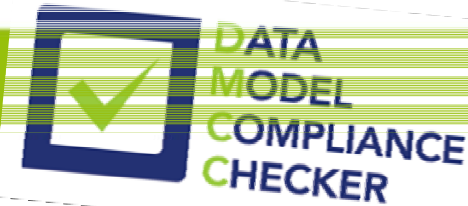
# Application Flowchart



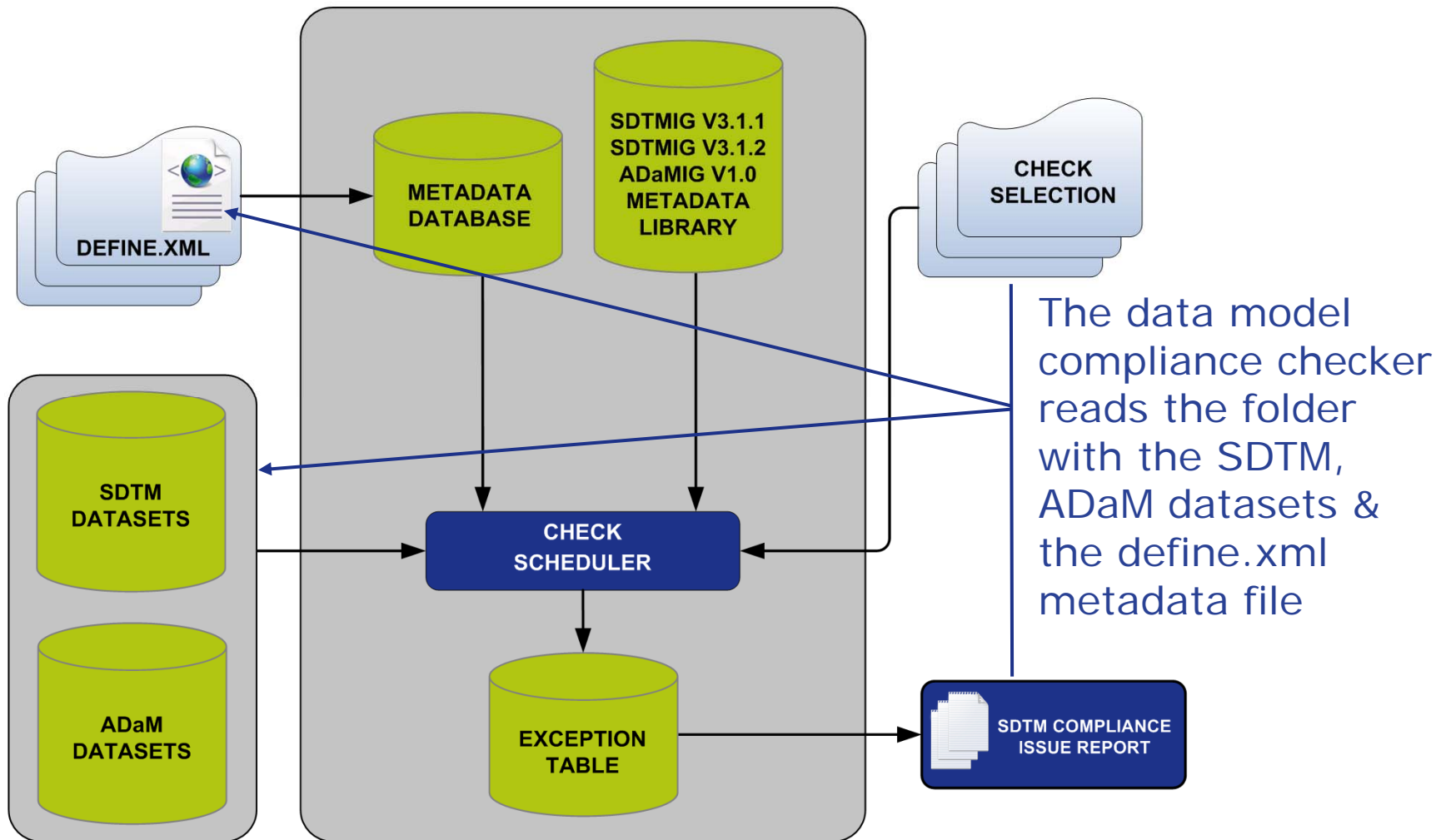
## DATA & METADATA ENVIRONMENT



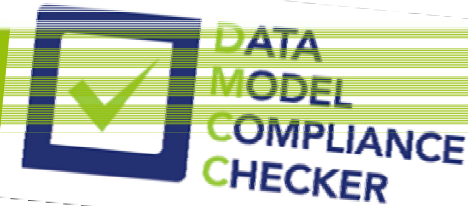
# Application Flowchart



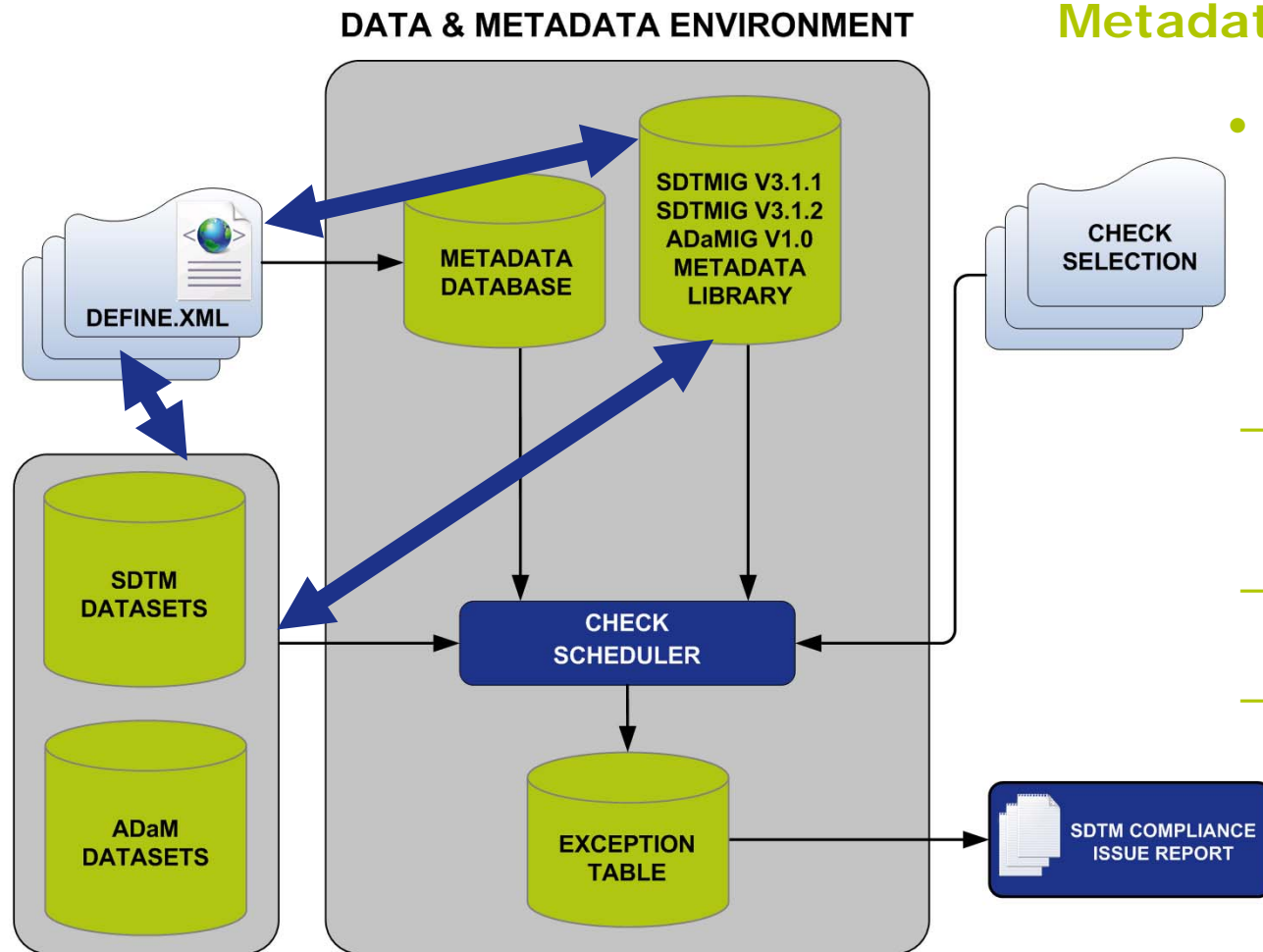
## DATA & METADATA ENVIRONMENT



# Application Flowchart

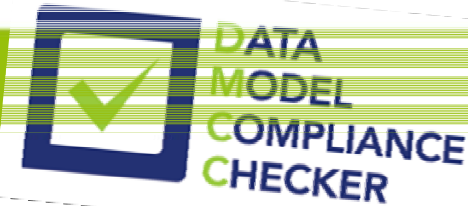


## Metadata Cross Checks

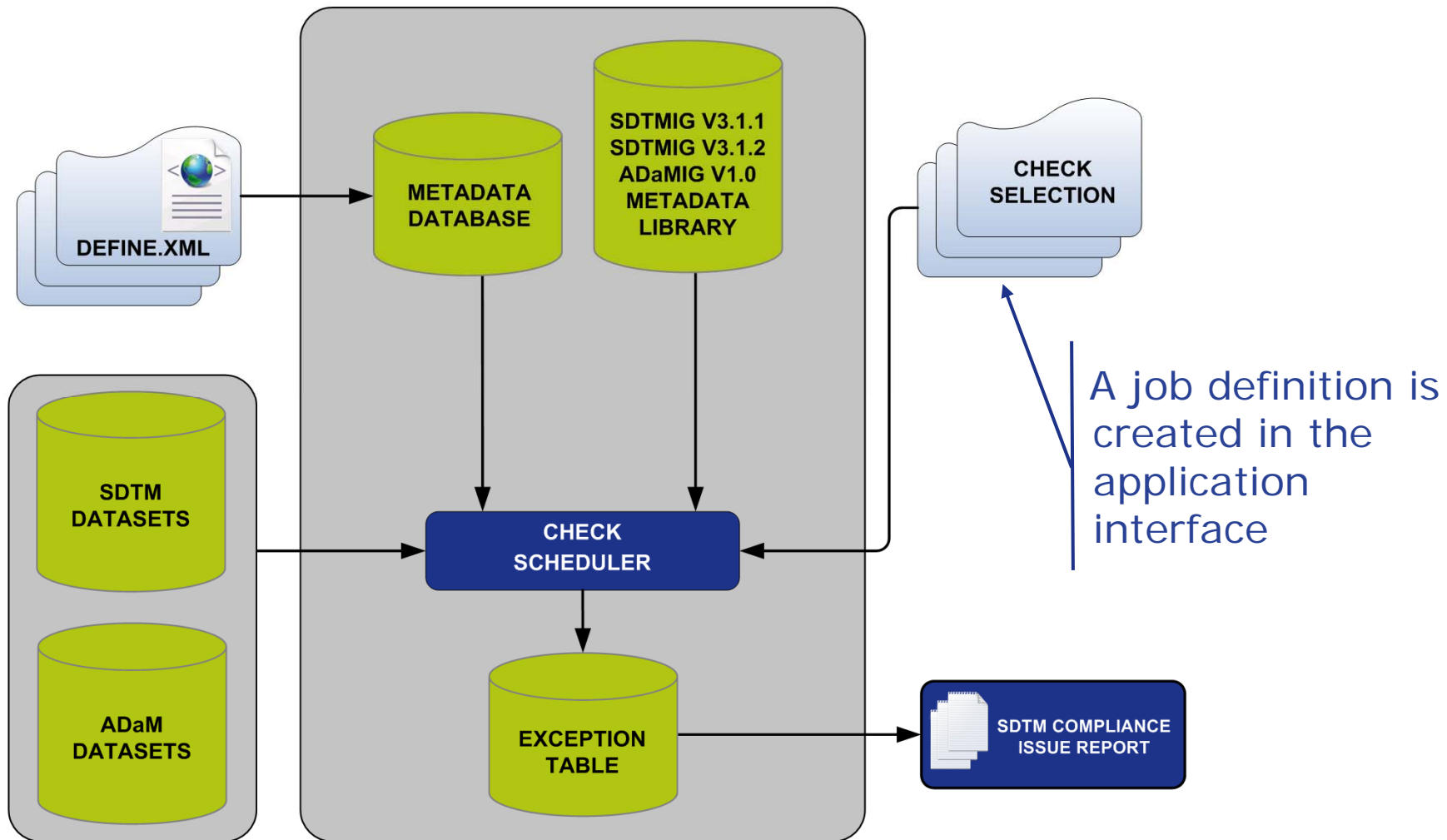


- Data Model Compliance Checks cross-check for consistency between:
  - CDISC SDTM or ADaM datasets definition
  - Metadata reported in the Define.xml
  - CDISC SDTM, ADaM or sponsor proprietary data standards in the metadata library

# Application Flowchart

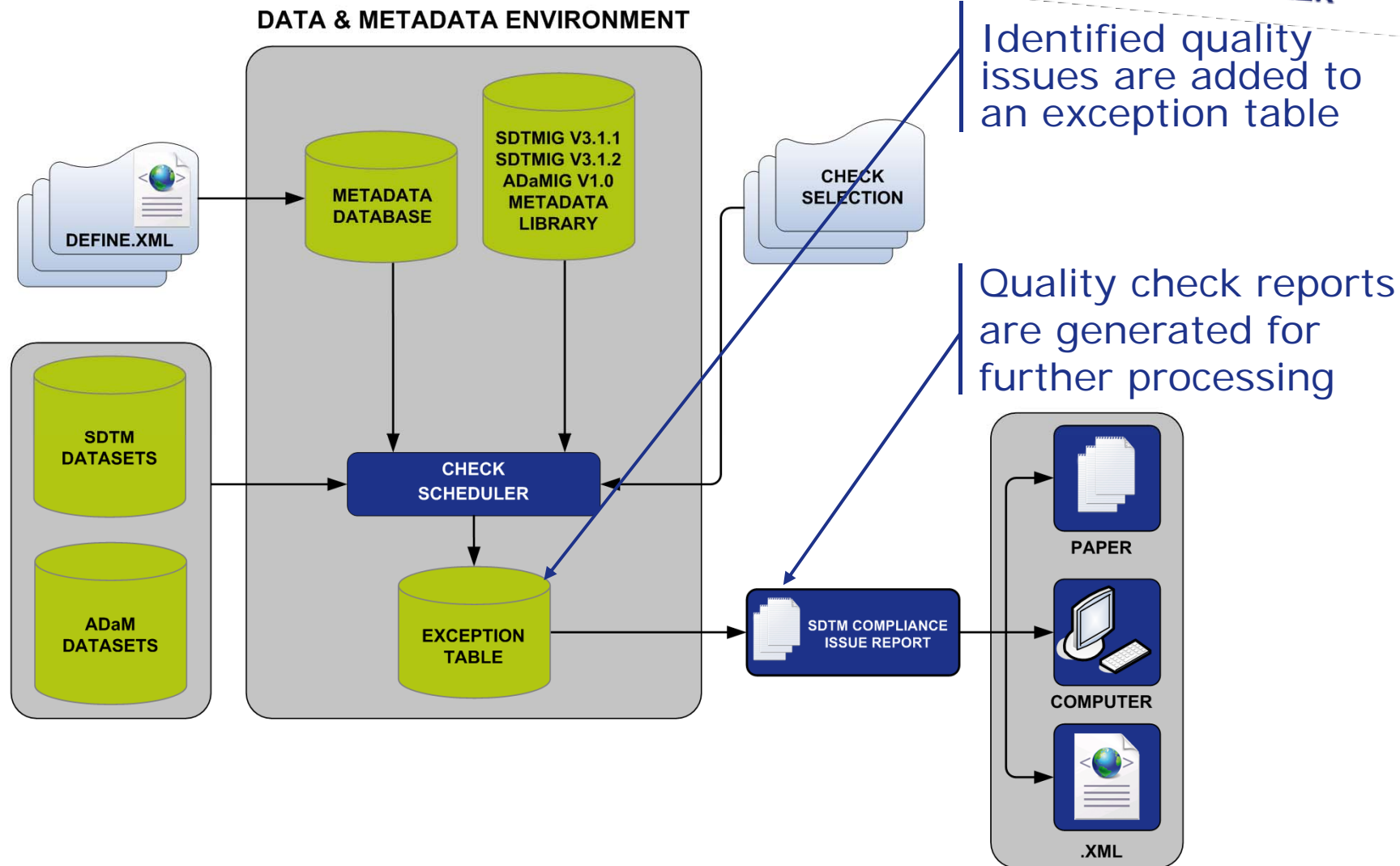
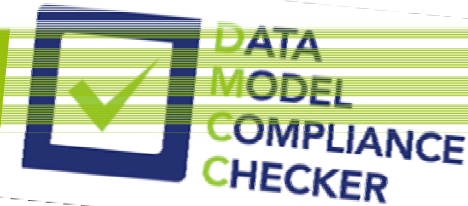


## DATA & METADATA ENVIRONMENT

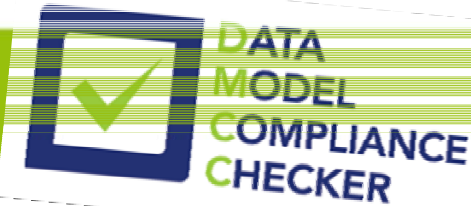




# Application Flowchart



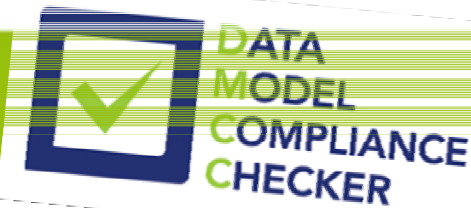




## Reporting Tool (1)

- 3 reports are created:
  - **Exception log**
  - Exception summary
  - Exception table
- > 50 exceptions per check per domain: only 3 exceptions are reported

LIST_OF_CHECKS	ERROR_LEVEL	COMPLETION_TIME	ELAPSED_TIME
Check1001	The Check Completed Successfully	2010-02-12T18-57-58	.
Check1002	The Check Completed Successfully	2010-02-12T18-57-58	0:00:00
Check1003	The Check Completed Successfully	2010-02-12T18-57-59	0:00:01
Check1004	The Check Completed Successfully	2010-02-12T18-57-59	0:00:00
Check1005	The Check Completed Successfully	2010-02-12T18-58-00	0:00:01
Check1006	The Check Completed Successfully	2010-02-12T18-58-00	0:00:00
Check1007	The Check Completed Successfully	2010-02-12T18-58-01	0:00:01
Check1008	The Check Completed Successfully	2010-02-12T18-58-02	0:00:01
Check1009	The Check Completed Successfully	2010-02-12T18-58-03	0:00:01
Check1010	The Check Completed Successfully	2010-02-12T18-58-19	0:00:16
Check1011	The Check Completed Successfully	2010-02-12T18-58-20	0:00:01
Check1012	The Check Completed Successfully	2010-02-12T18-58-37	0:00:17
Check1013	The Check Completed Successfully	2010-02-12T18-58-37	0:00:00
Check1014	The Check Completed Successfully	2010-02-12T18-58-40	0:00:03
Check1015	The Check Completed Successfully	2010-02-12T18-58-41	0:00:01
Check1016	The Check Completed Successfully	2010-02-12T18-58-41	0:00:00
Check1017	The Check Completed Successfully	2010-02-12T18-58-43	0:00:02
Check1018	The Check Completed Successfully	2010-02-12T18-58-43	0:00:00
Check1019	The Check Completed Successfully	2010-02-12T18-58-44	0:00:01
Check1020	The Check Completed Successfully	2010-02-12T18-58-44	0:00:00
Check1021	The Check Completed Successfully	2010-02-12T18-58-44	0:00:00
Check1022	The Check Completed Successfully	2010-02-12T18-58-45	0:00:01
Check1023	The Check Completed Successfully	2010-02-12T18-58-45	0:00:00



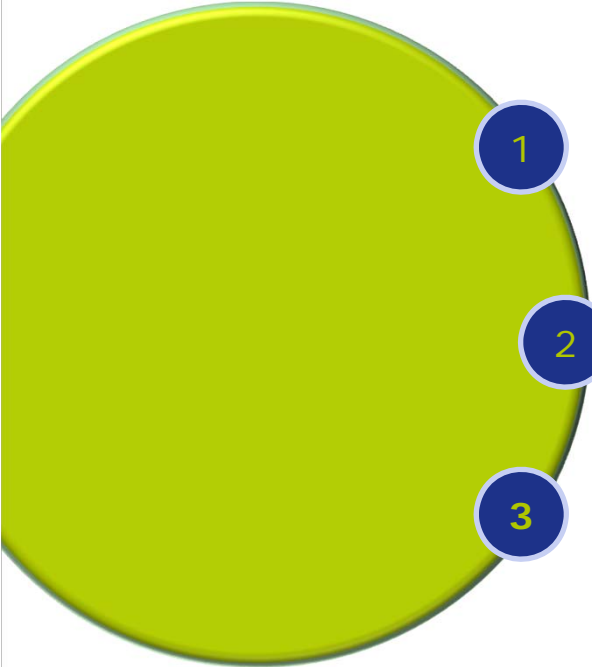
## Reporting Tool (2)

- 3 reports are created:
  - Exception log
  - **Exception summary**
  - Exception table
- > 50 exceptions per check per domain: only 3 exceptions are reported

LIST OF CHECKS	DOMAIN	VARIABLE1	EXCEPTIONS COUNT	CHECK TYPE
Check0009	LB	LBORRES	73	DATA
Check0010	LB	LBORRESU	84	METADATA
Check0010	VS	VSORRESU	56	METADATA
Check0010	VS	VSSTRESU	56	METADATA
Check0015	CO	IDVAR	73	DATA
Check0100	IE	DOMAIN	1	DATA
Check0268	CM	COLUMN	1	METADATA
Check0271	VS	COMPALGO	5	METADATA



## Topics

- 
- 1 Current Situation
  - 2 Applied CDISC Compliance
  - 3 Summary

## Summary

- CDISC compliance will facilitate:
  - The use of the FDA tools
  - In-house communication between CRO and sponsors
- Running these checks do not necessarily solve all findings but sponsor is prepared for FDA feedback
  - Document confirmed issues in FDA Reviewers' Guide
  - CBER calls this the 'Validation and Data Interpretation Report'
- FDA will not reject your submission (yet)!



LIFE SCIENCES

**Thank you  
for your attention**

**Nick De Donder  
Project Manager  
Business & Decision Life Sciences**

Tel +32 2 774 11 00  
Fax +32 2 774 11 99  
Mobile +32 476 82 20 27  
[nick.dedonder@businessdecision.com](mailto:nick.dedonder@businessdecision.com)

Sint-Lambertusstraat 141 Rue Saint-  
Lambert  
1200 Brussels

[www.businessdecision-lifesciences.com](http://www.businessdecision-lifesciences.com)



**DRIVE YOUR PERFORMANCE®**