

Towards 360i, or at least 180i: A Validation Perspective

Marcelina Hungria and Allan Hungria, Dilcore Group

Date: 13-Oct-2025





Towards 360i, or at least 180i: A Validation Perspective

Presented by:

Marcelina Hungria, Dlcore Group, Managing Director / Consultant
Allan Hungria, Dlcore Group, Consultant

Meet the Speakers

Marcelina Hungria, MBA

Title: Managing Director / Consultant

Organization: Dlcore Group

Marcelina Hungria is a Computer Science graduate, and MBA from Rutgers University. Marcelina has been supporting the pharma/biotech industry with Metadata Repository (MDR) development, CDISC Standards Implementation and Data Submission services for nearly 30 years. She has extensive experience developing and validating software tools, plus overall system testing, using SAS and other programming languages.

CDISC volunteer for 17 years, actively participating on the CDISC Data-Exchange, SDTM and SDTMIG development teams, plus the CDISC CORE initiative. Marcelina is a CDISC authorized Instructor for CDISC Data-Exchange standards.

Allan Hungria, Ph. D.

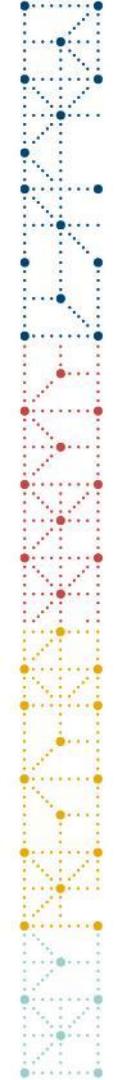
Title: Consultant

Organization: Dlcore Group

Allan Hungria graduated with his doctorate in applied mathematics from the University of Delaware in 2019. His professional interests lie somewhere between numerical computation, programming and teaching. Since 2019, he has worked as a consultant, a data scientist using MATLAB and Mathematica, python, R, PostgreSQL among others, and a professor of mathematics and statistics.

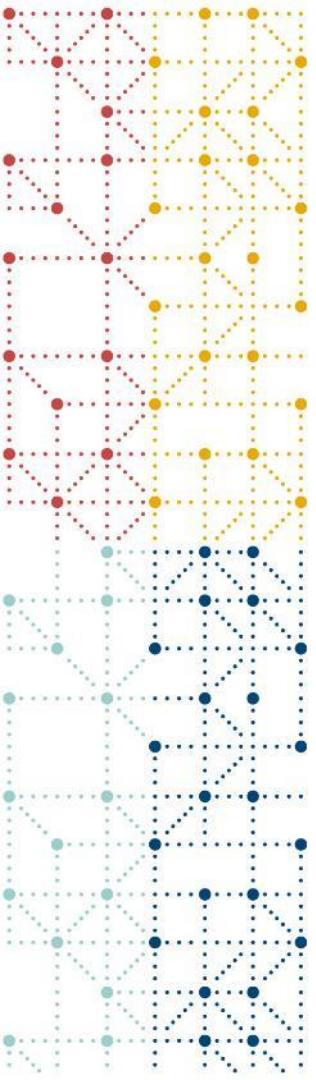
Related to CDISC, Allan has participated in diverse projects for processing of data standards CDASH, SDTM, ADaM and Controlled Terminology, in ODM representation via the CDISC Library API, replication of TLFs for cross-validation of submission data packages, and co-developing Dlcore Group's software tools including the Dataset-JSON Viewer and Dataset-JSON Validation.





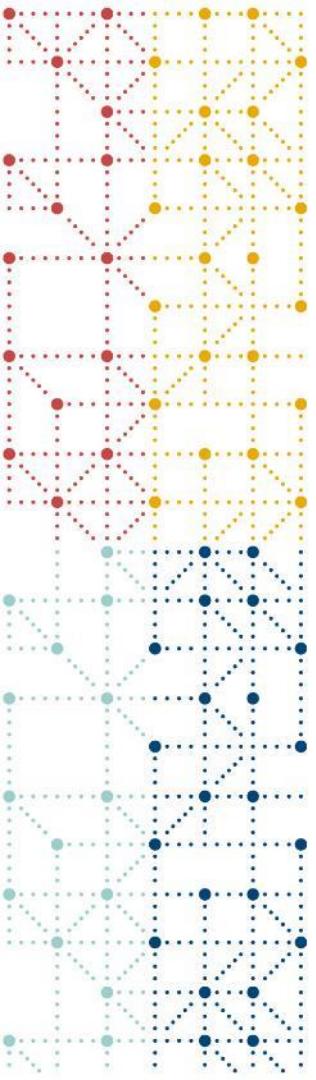
Disclaimer and Disclosures

- *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.*
- *{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. A high-level glance at CDISC 360
2. 360i: Data exchange Process, Dataset-JSON API - demo
3. 360i: Dataset-JSON Visualization, Integrating Validation Tools in the process - demo
4. 360i Validation: Summary of Challenges and forward steps



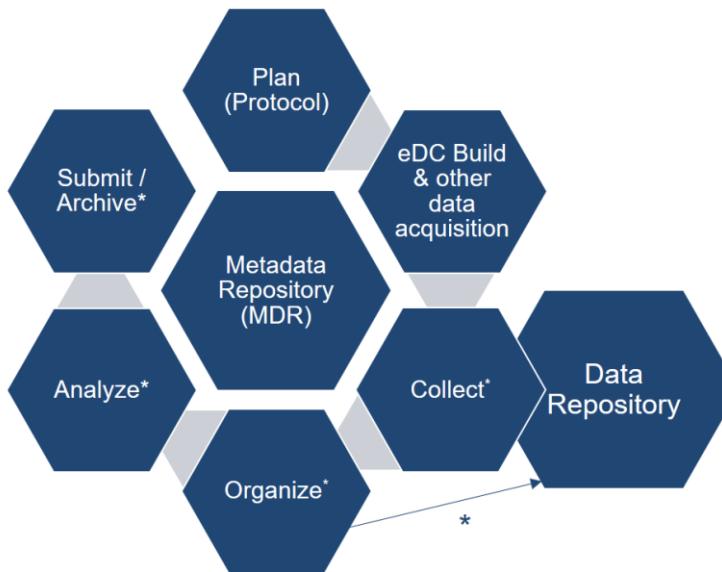
A high-level glance at CDISC 360

Process and Standards

Research Process

Silos vs Integrated approach:

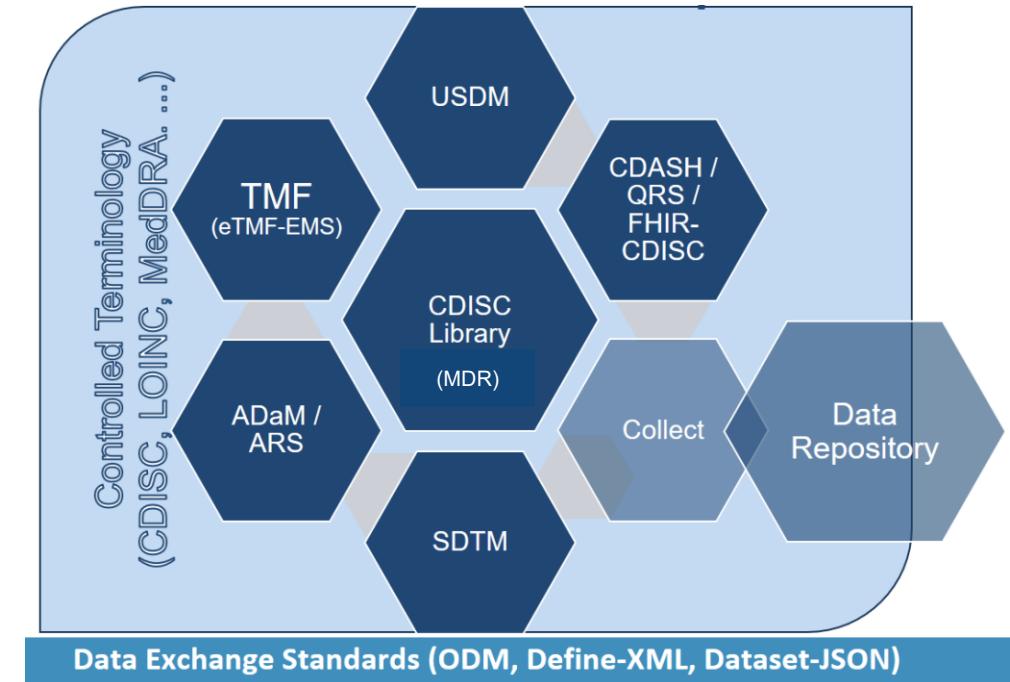
- Process
- Systems



Research Process CDISC Standards Implementation

Silos vs Integrated approach:

- Disparate systems is a reality.



Research Process Metadata and Data repositories

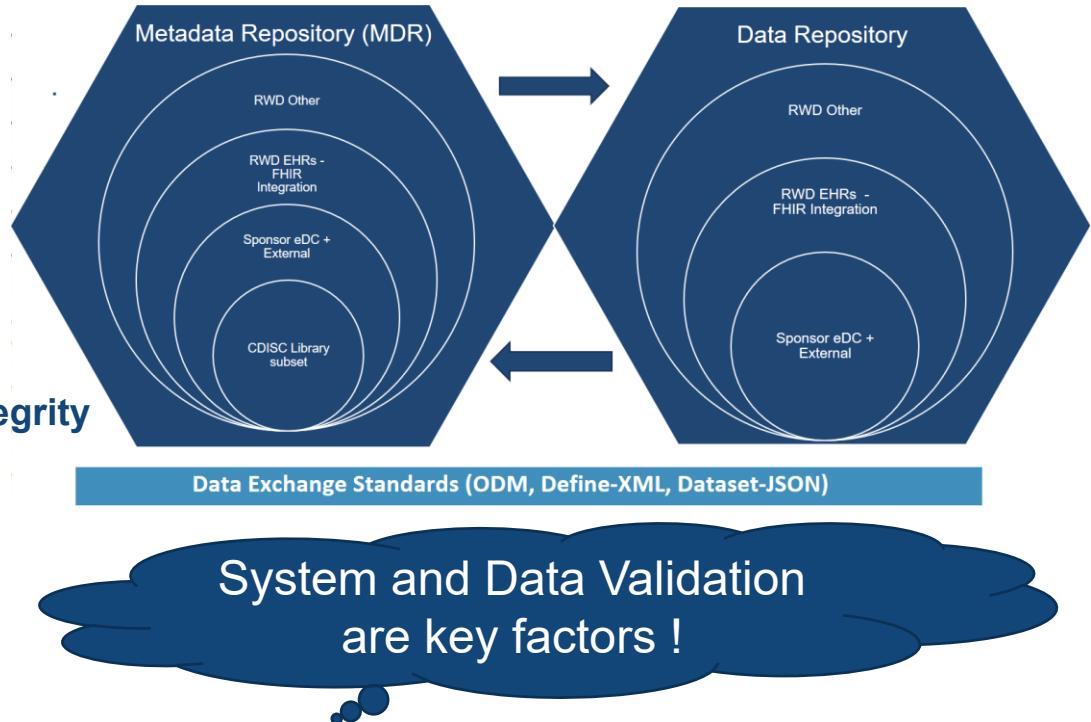
Underlying Infrastructure

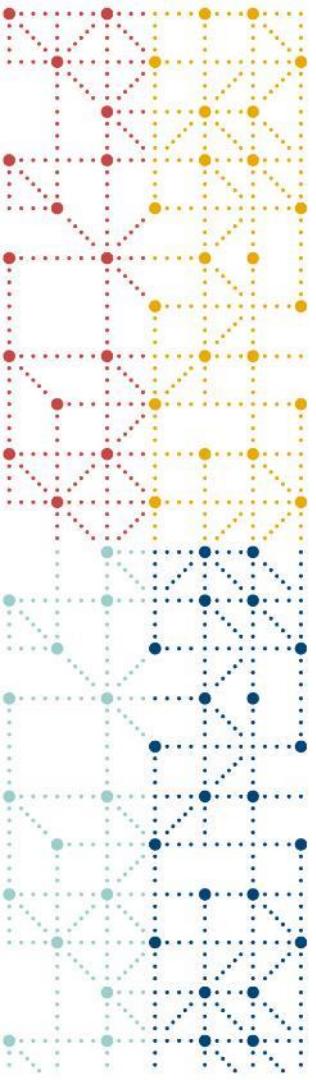
- Different Operating systems
- Different programming languages
- Different Software Applications
 - Open source vs proprietary
 - Freely available vs licensed
 - Open source challenges
- Fast-pace technology changes

Focus on Interoperability & Data integrity

Validation:

- Across different systems
- Across different organizations
- Validation at different levels





360i: Data Exchange Process

Summary

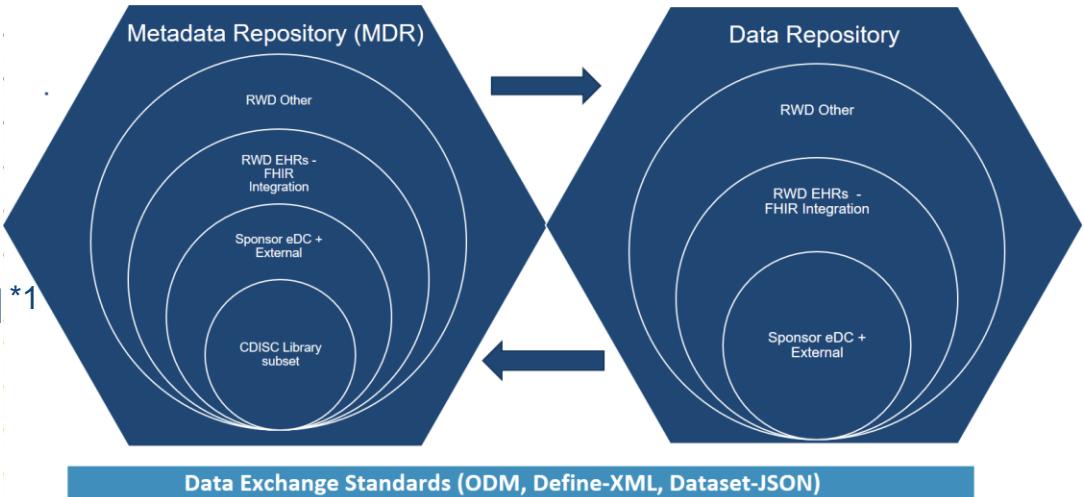
Data Exchange Process

Metadata

- Hierarchical: ODMv2^{*1} ^{*2}
- Tabular: Define-XML^{*2}

Data

- Tabular: Dataset-JSON v1.1^{*1}
- Hierarchical: ODMv2^{*1} ^{*2}



*1 Encouraging adoption of newest CDISC standards

*2 JSON representation is WIP

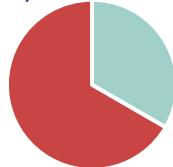
System and Data Validation
are key factors !

Data-Exchange Standards Adoption

Standards

- ODM
 - V2 (released Aug-2023)
 - V1.3.2 (released Dec-2013)
- Define-XML (ODMv1.3.2 based)^{*}
 - V2.1 (released Jan-2019)
 - V2.0 (released Mar-2013)
- Dataset-JSON
 - V1.1 (released Dec-2024)
 - [FDA requested comments](#)
 - [Dlcore Group's comments](#)
 - Dataset-JSON API
 - V1.0 ([Public Review](#) ends 17-Oct-2025)

• [CDISC ODM Certified Products](#)
(mostly EDC Systems)



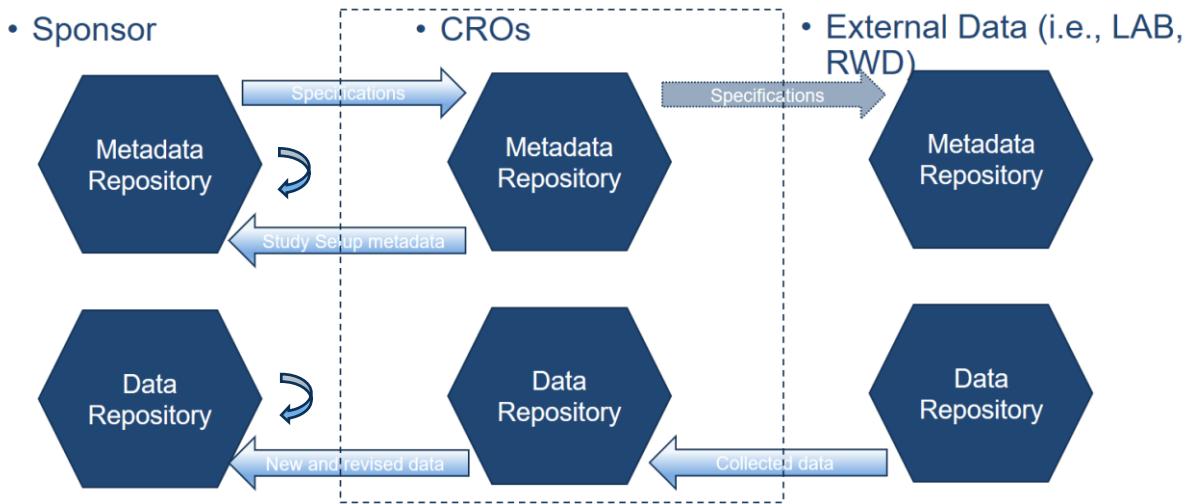
■ ODM v2 ■ ODM v1.3.2 ■ ODM v1.3.1 and older

- No statistics publicly available. V2.1 usage increasing based on comments received during training
- At least 10 [Dataset-JSON viewers](#) available to the user community
- Dataset-JSON API Reference implementation Incorporated in the [Dlcore Group set of tools](#) and other Dataset-JSON viewers.

Data-Exchange Process: Acquisition or Transformation

Data Integrity & Validation:

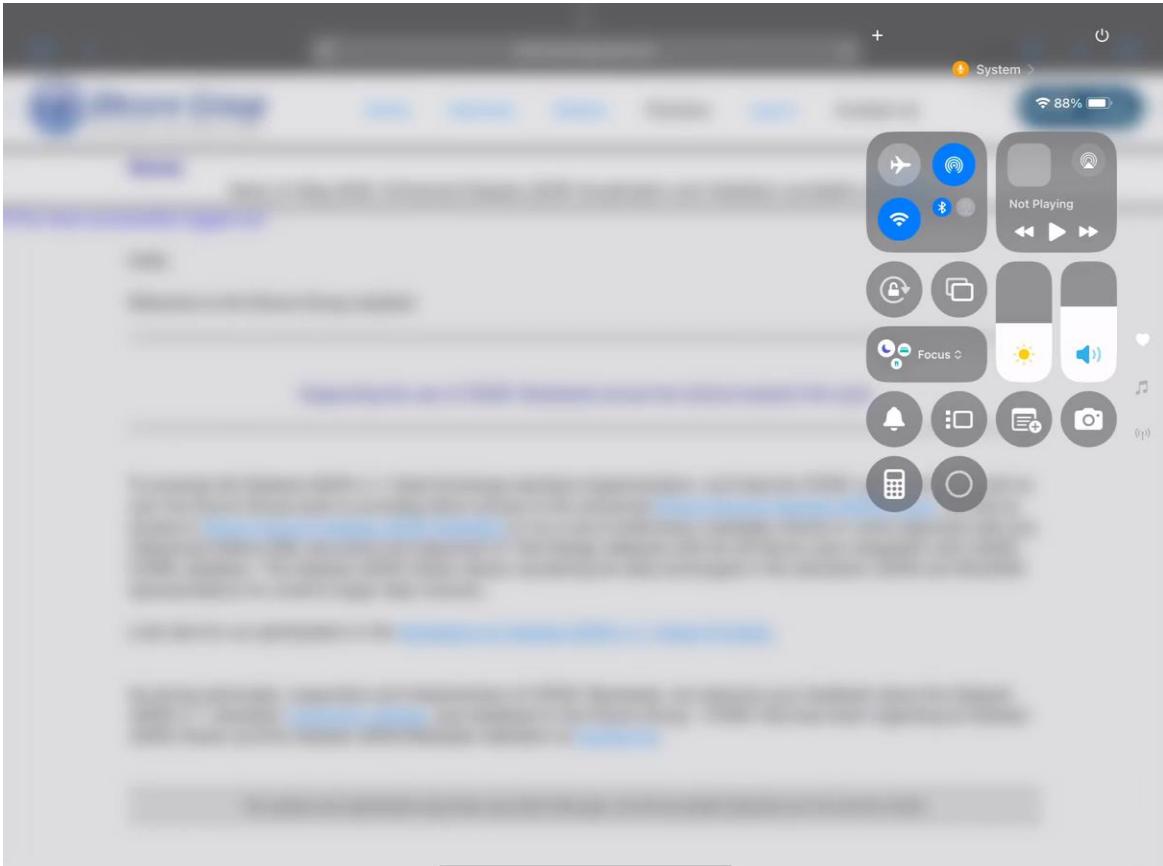
- Specification ~ Data Transfer Agreement (DTA)



Data Exchange Process – Acquisition

Acquiring data

- APIs (cache)
- Other mechanisms
 - i.e.; sftp



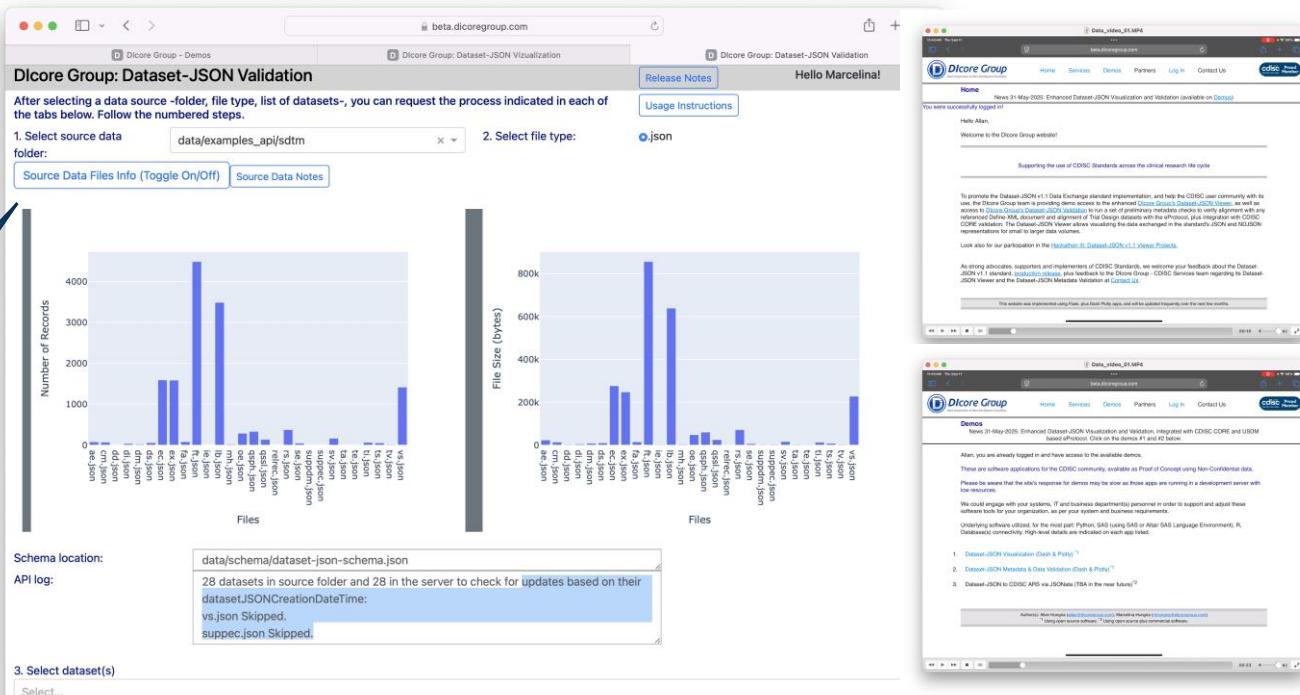
Data Exchange Process – Acquisition

Acquiring data

- APIs^{*1} (cache)
- Other mechanisms
 - i.e.: sftp

DTA

*1 Reference Implementation



Data Exchange Process – Acquisition & Validation

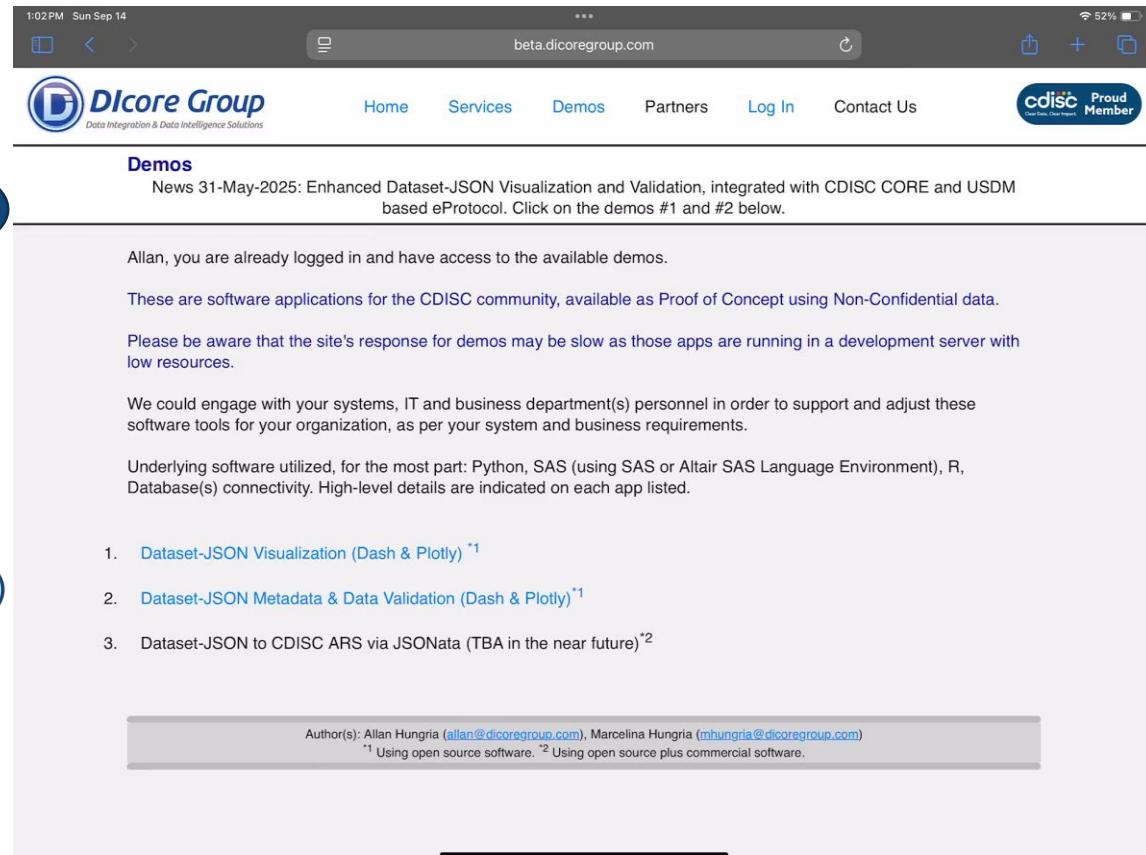
Acquiring data

- APIs (cache)
- Other mechanisms

DTA

Validating data (cache)

- Structural Validation
 - Schema
 - Other
- High-level inspection
 - Source Dataset-JSON
 - Alignment with Define-XML
- Data Validation (CORE Engine)
 - Standard Conformance
 - Custom Rules groups
- Data against e-Protocol
- Data against e-SAP



The screenshot shows a mobile browser displaying the beta.dicoregroup.com website. The header includes the Dicore Group logo, navigation links for Home, Services, Demos, Partners, Log In, and Contact Us, and a CDISC Proud Member badge. The main content area is titled "Demos" and features a news item: "News 31-May-2025: Enhanced Dataset-JSON Visualization and Validation, integrated with CDISC CORE and USDM based eProtocol. Click on the demos #1 and #2 below." Below this, a message says "Allan, you are already logged in and have access to the available demos." It then describes the demos as "software applications for the CDISC community, available as Proof of Concept using Non-Confidential data." It cautions that site response may be slow due to development server resources. It also notes that engagement with IT and business departments is required for support. The underlying software is listed as Python, SAS (using SAS or Altair SAS Language Environment), R, and Database(s) connectivity. A numbered list of three demos is provided, with the first two marked as using open source software and the third as using open source plus commercial software. The footer contains author information and a note about using open source software.

1. Dataset-JSON Visualization (Dash & Plotly)^{*1}

2. Dataset-JSON Metadata & Data Validation (Dash & Plotly)^{*1}

3. Dataset-JSON to CDISC ARS via JSONata (TBA in the near future)^{*2}

Data Exchange Process – Acquisition & Validation

Acquiring data

- APIs (cache)
- Other mechanisms

Validating data (cache)

- Structural Validation
 - Schema
 - Other
- High-level inspection
 - Source Dataset-JSON
 - Alignment with Define-XML
- Data Validation ([CORE Engine](#)^{*1})
 - Standard Conformance
 - Custom Rules groups
- Data against e-Protocol
- Data against e-SAP

*1 Reference Implementation

The screenshot displays the Dicore Group Dataset-JSON Validation and Visualization interface. The validation tab shows a JSON schema, a JSON tree, and a feedback section. The visualization tab displays two bar charts showing the number of records and file size for various datasets. The validation tab for the CORE Engine shows a log and report with validation results.

Dicore Group: Dataset-JSON Validation

After selecting a data source -folder, file type, list of datasets-, you can request the process indicated in each of the tabs below. Follow the numbered steps.

1. Select source data
2. Select file type:
3. Select dataset(s)
4. Click on each tab below to process your request

Feedback

Please help us to improve the Dataset-JSON Validation by rating its overall functionality and usability on the scale below.

Dicore Group: Dataset-JSON Visualization

After selecting a data source -folder, file type, list of datasets-, you can request the process indicated in each of the tabs below. Follow the numbered steps.

1. Select source data
2. Select file type:

Source Data Files Info (Toggle On/Off) **Source Data Notes**

Schema location:
API log: 28 datasets in source folder and 28 in the server to check for updates based on their datasetJSONCreationDateTime.

Number of Records

Files

File Size (bytes)

Schema Validation **Display JSON file** **Alignment with Define-XML** **Alignment with eProtocol** **Data Validation (CDISC CORE)** **Reset tabs**

CORE Validation log

```
CORE Validation Report already exists core_val_data_e
CORE Validation
source /Users/marcelinehungria/Documents/WIX01/cor
[Level 5] 2025-09-11 13:09:26,870 - console_logger.py
[Level 5] 2025-09-11 13:09:27,086 - console_logger.py
[Level 5] 2025-09-11 13:09:27,105 - console_logger.py
[Level 5] 2025-09-11 13:09:27,120 - console_logger.py
[Level 5] 2025-09-11 13:09:27,321 - console_logger.py
[Level 5] 2025-09-11 13:09:27,502 - console_logger.py
[Level 5] 2025-09-11 13:09:27,758 - console_logger.py
[Level 5] 2025-09-11 13:09:27,875 - console_logger.py
[Level 5] 2025-09-11 13:09:27,923 - console_logger.py
```

CORE Validation report

```
core_id : CORE-80089
message : LBORRES is not a continuous measurement or Empty, when LBORRH1 is not empty.
executability : fully executable
dataset : 1b.json
USUBID : CDISC001
row : 87
```

Feedback

Please help us to improve the Dataset-JSON Validation by rating its overall functionality and usability on the scale below.

To: cdiscservices@dicoregroup.com

Data Exchange Process – Acquisition & Validation

Acquiring data

- APIs (cache)
- Other mechanisms

Validating data (cache)

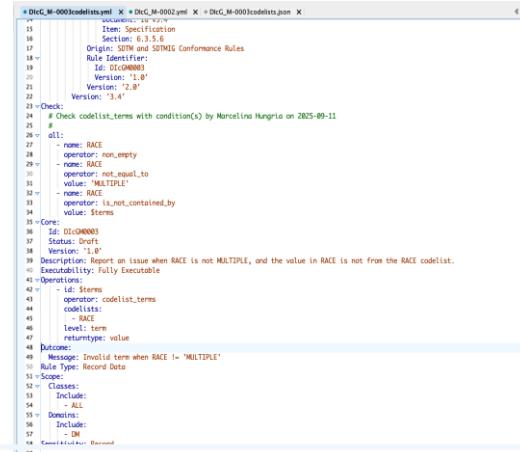
- Structural Validation
 - Schema
 - Other
- High-level inspection
 - Source Dataset-JSON
 - Alignment with Define-XML
- Data Validation (CORE Engine)
 - Standard Conformance
 - Regulatory Rules
 - Additional Rules
- Custom Rules groups
- Data against e-Protocol
- Data against e-SAP

Group n (for DTA xyz)

Sponsor / Others (i.e. Sxyz v# & DlcG v1.0 & ...)

Regulatory (FDA x or PMDA y or ...)

Standard Conformance (i.e. (SDTMIG 3.4 & SDTMIGMD 1.0), (ADaMIG 1.3 & related), USDM v4.0), ...

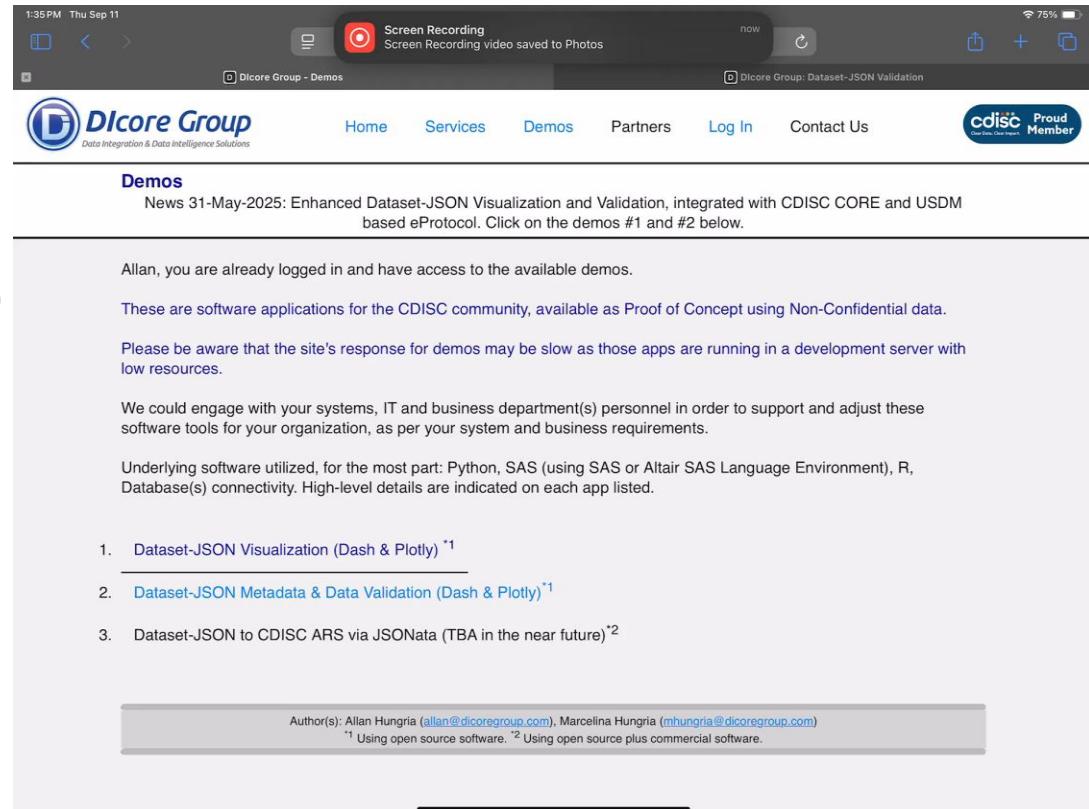


```
 1  # DlcG_M-000/codelists.xml  X  = DlcG_M-0002.xml  X  = DlcG_M-000/codelists.json  X
 2
 3  Item Specification
 4  Section: 6.3.5.6
 5  Original Source: SDTMIG Conformance Rules
 6  Rule Identifier:
 7  Id: DlcG0000
 8  Version: 1.0
 9  Status: Draft
10  Version: 1.0
11  Version: 1.4
12
13  <!--Check codelist_terms with condition() by Marcelina Hungria on 2025-09-11
14  # -->
15  <!--
16  <!--
17  <!--
18  <!--
19  <!--
20  <!--
21  <!--
22  <!--
23  <!--
24  <!--
25  <!--
26  <!--
27  <!--
28  <!--
29  <!--
30  <!--
31  <!--
32  <!--
33  <!--
34  <!--
35  <!--
36  <!--
37  <!--
38  <!--
39  <!--
40  <!--
41  <!--
42  <!--
43  <!--
44  <!--
45  <!--
46  <!--
47  <!--
48  <!--
49  <!--
50  <!--
51  <!--
52  <!--
53  <!--
54  <!--
55  <!--
56  <!--
57  <!--
58  <!--
59  <!--
```

Data Exchange Process –Visualization

After Structural Validation

- High-level inspection
 - Metadata
 - Data summary / statistics (cache)
 - Data details
 - Default view
 - data sample for large volumes
 - Selective view
 - Data streaming
- Highlighted validation issues
 - At a data item level
 - i.e., DM



1:35PM Thu Sep 11

Screen Recording
Screen Recording video saved to Photos

Dicore Group - Demos

Dicore Group - Dataset-JSON Validation

Dicore Group
Data Integration & Data Intelligence Solutions

Home Services Demos Partners Log In Contact Us

Demos

News 31-May-2025: Enhanced Dataset-JSON Visualization and Validation, integrated with CDISC CORE and USDM based eProtocol. Click on the demos #1 and #2 below.

Allan, you are already logged in and have access to the available demos.

These are software applications for the CDISC community, available as Proof of Concept using Non-Confidential data.

Please be aware that the site's response for demos may be slow as those apps are running in a development server with low resources.

We could engage with your systems, IT and business department(s) personnel in order to support and adjust these software tools for your organization, as per your system and business requirements.

Underlying software utilized, for the most part: Python, SAS (using SAS or Altair SAS Language Environment), R, Database(s) connectivity. High-level details are indicated on each app listed.

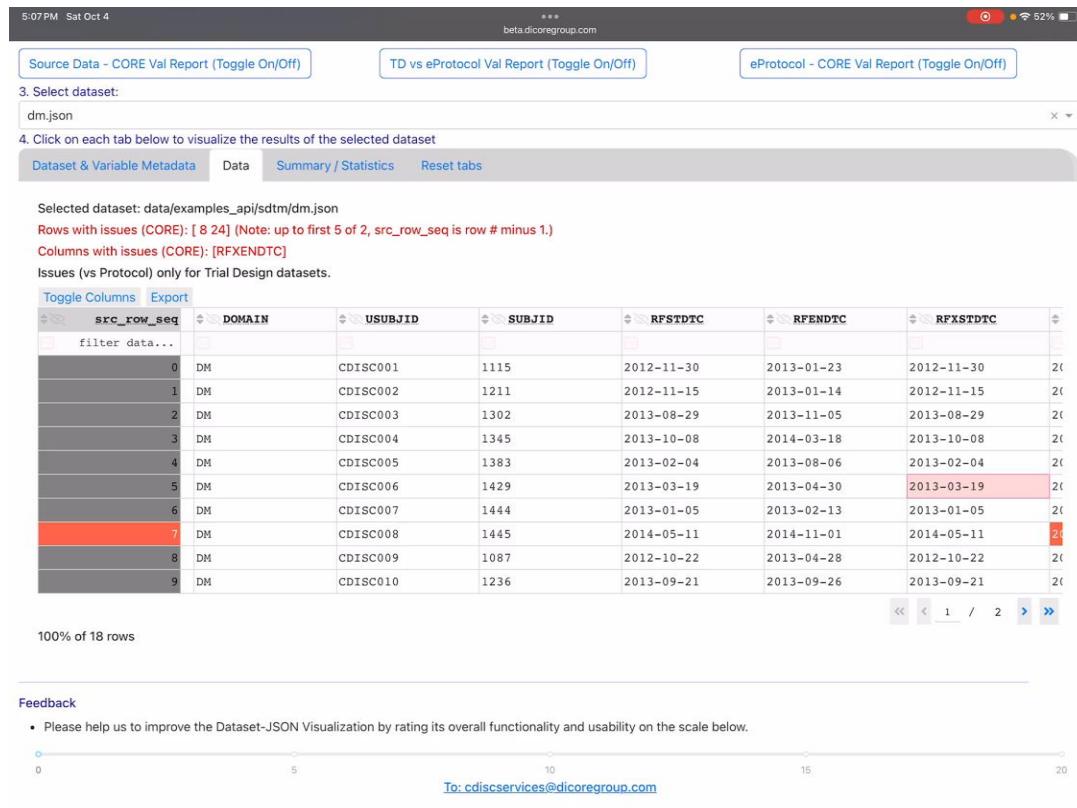
1. Dataset-JSON Visualization (Dash & Plotly) ^{*1}
2. Dataset-JSON Metadata & Data Validation (Dash & Plotly)^{*1}
3. Dataset-JSON to CDISC ARS via JSONata (TBA in the near future)^{*2}

Author(s): Allan Hungria (allan@dicoregroup.com), Marcelina Hungria (mhungria@dicoregroup.com)
*1 Using open source software. *2 Using open source plus commercial software.

Data Exchange Process –Visualization, cont.

After Structural Validation

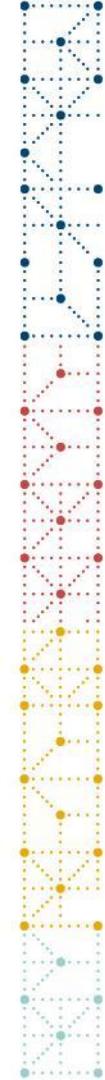
- High-level inspection
 - Metadata
 - Data summary / statistics (cache)
 - Data details
 - Default view
 - data sample for large volumes
 - Selective view
 - Data streaming
- Highlighted validation issues
 - At a data item level
 - i.e., LB



The screenshot shows a web-based data visualization tool for validating datasets. At the top, there are three tabs: 'Source Data - CORE Val Report (Toggle On/Off)', 'TD vs eProtocol Val Report (Toggle On/Off)', and 'eProtocol - CORE Val Report (Toggle On/Off)'. Below the tabs, a message says '3. Select dataset: dm.json'. Another message indicates '4. Click on each tab below to visualize the results of the selected dataset'. The 'Dataset & Variable Metadata' tab is selected, showing 'Summary / Statistics' and 'Reset tabs' options. A note says 'Selected dataset: data/examples_api/sdtm/dm.json'. It also lists 'Rows with issues (CORE): [8 24] (Note: up to first 5 of 2, src_row_seq is row # minus 1.)' and 'Columns with issues (CORE): [RFXENDTC]'. A note states 'Issues (vs Protocol) only for Trial Design datasets.' The main area is a table with the following data:

src_row_seq	DOMAIN	USUBJID	SUBJID	RFSTDTC	RFENDTC	RFXSTDTC
filter data...						
0	DM	CDISC001	1115	2012-11-30	2013-01-23	2012-11-30
1	DM	CDISC002	1211	2012-11-15	2013-01-14	2012-11-15
2	DM	CDISC003	1302	2013-08-29	2013-11-05	2013-08-29
3	DM	CDISC004	1345	2013-10-08	2014-03-18	2013-10-08
4	DM	CDISC005	1383	2013-02-04	2013-08-06	2013-02-04
5	DM	CDISC006	1429	2013-03-19	2013-04-30	2013-03-19
6	DM	CDISC007	1444	2013-01-05	2013-02-13	2013-01-05
7	DM	CDISC008	1445	2014-05-11	2014-11-01	2014-05-11
8	DM	CDISC009	1087	2012-10-22	2013-04-28	2012-10-22
9	DM	CDISC010	1236	2013-09-21	2013-09-26	2013-09-21

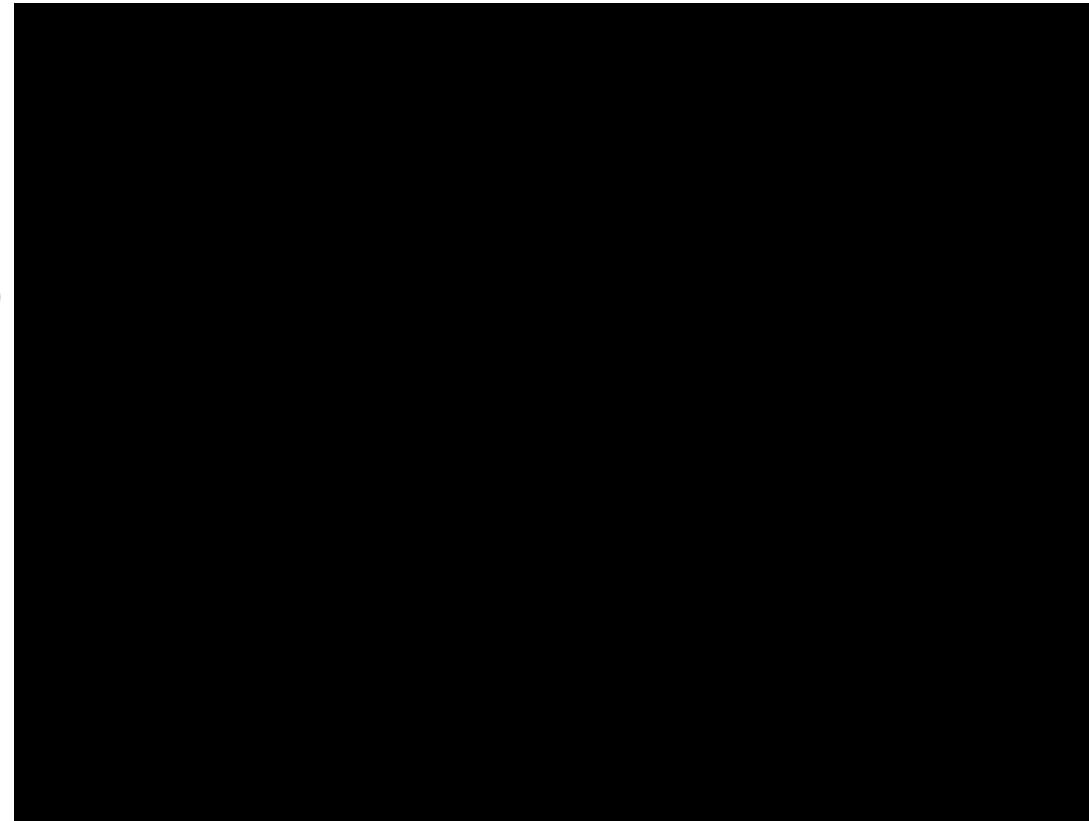
Below the table, a note says '100% of 18 rows'. A 'Feedback' section asks for ratings on a scale from 0 to 20, with a note to 'To: cdiscservices@dicoregroup.com'.

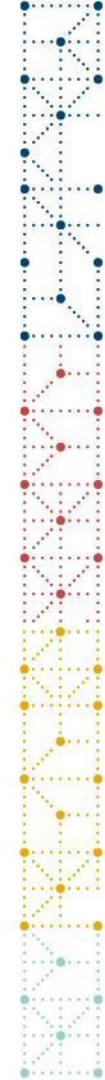


Data Exchange Process –Visualization, cont.

After Structural Validation

- High-level inspection
 - Metadata
 - Data summary / statistics (cache)
 - Data details
 - Default view
 - data sample for large volumes
 - Selective view
 - Data streaming
- Highlighted validation issues
 - At a data item level
 - i.e., TV (against the e-Protocol)



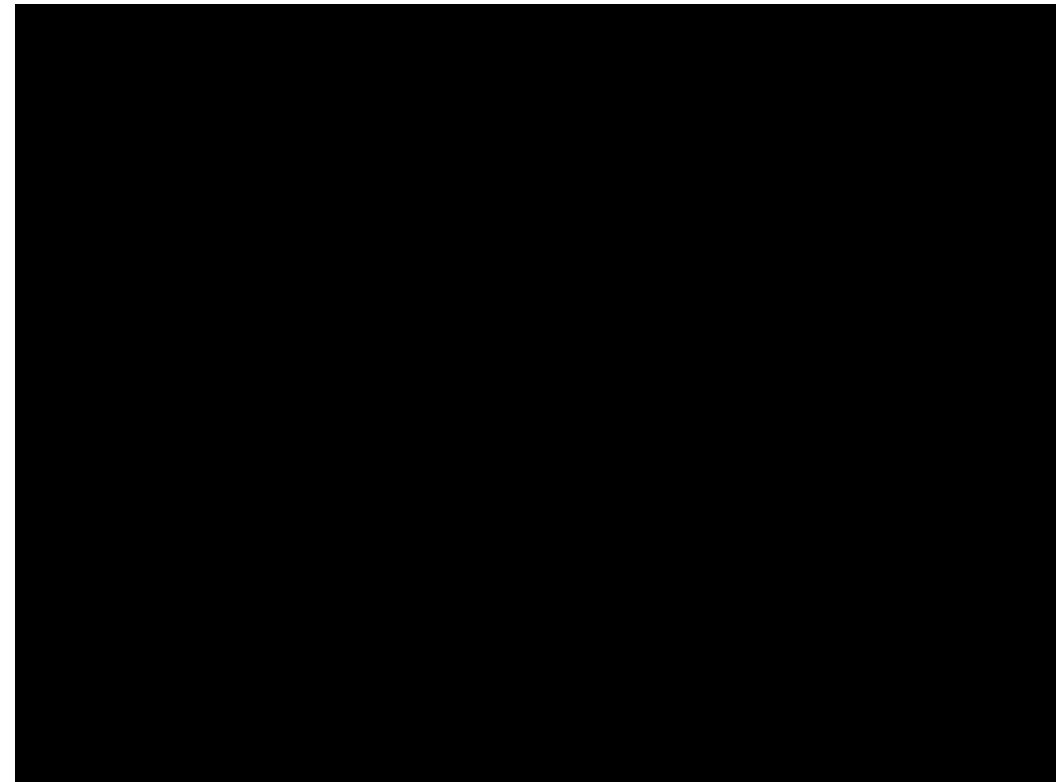


Data Exchange Process –Visualization, cont..

After Structural Validation

- High-level inspection
 - Metadata
 - Data summary / statistics (cache)
 - Data details
 - Default view
 - data sample for large volumes
 - Selective view
 - Data streaming
- Highlighted validation issues
 - At a data item level

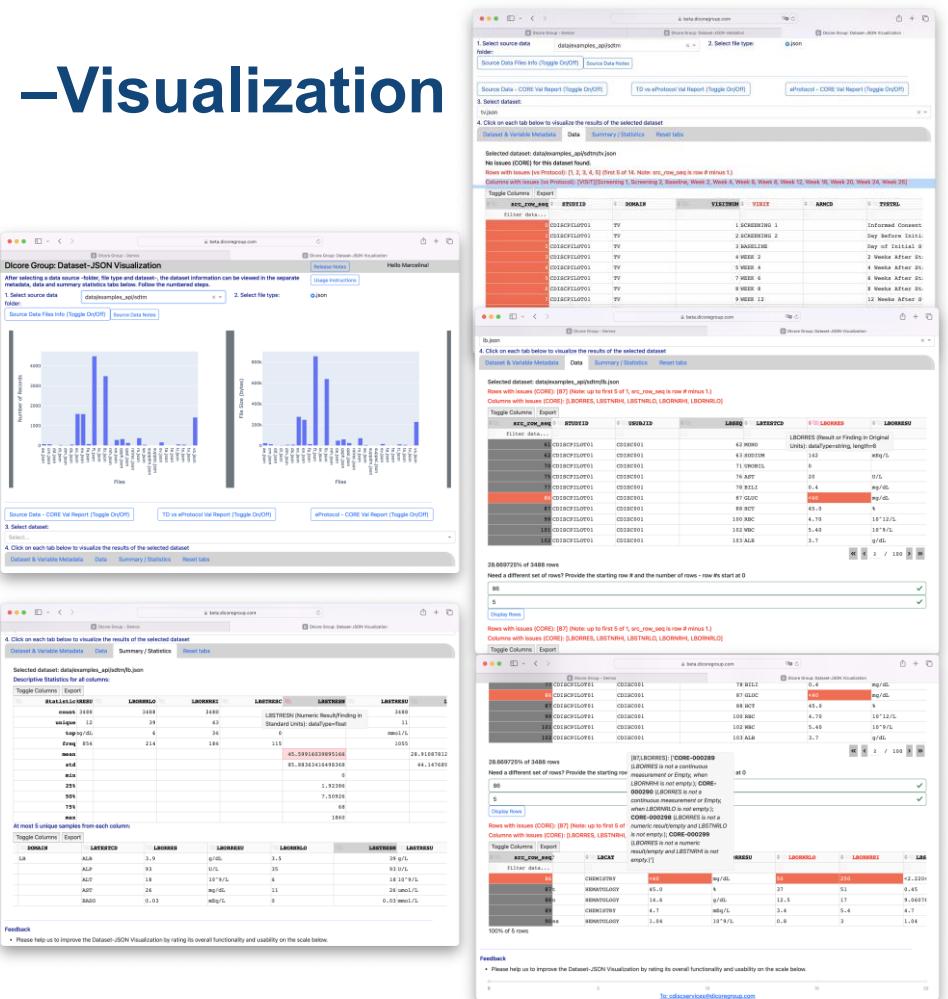
Caviat



Data Exchange Process –Visualization

After Structural Validation

- High-level inspection
 - Metadata
 - Data summary / statistics (cache)
 - Data details
 - Default view
 - data sample for large volumes
 - Selective view
 - Data streaming
- Highlighted validation issues
 - At a data item level
 - Reported using the CORE engine
 - Reported using different mechanism



#ClearDataClearImpact

360i Validation: Summary of Challenges and Forward Steps

Process and Tools

Fact: there are different groups in each organization and different software tools available to the user community.

- Determining and prioritizing validation points
 - Different use cases
- Understanding different standards.
 - Different schemas
 - Different versions
- Parsing different sources and outputs
 - Different systems
 - Different programming languages
- CDISC initiatives evolving; i.e., CDISC CORE
 - Engine re-work undergoing
 - Improvements to implementation and usability of rules; i.e., capability to provide detailed location of issue(s) reported

Encourage building common ground for data exchange

- Create/support infrastructure where diverse software tools be dynamically and programmatically invoked and their output consumed
 - Take advantage of available resources
 - Willingness to test ideas
 - Managed / validated infrastructure
- Gradually incorporate more validation and integration points
 - i.e., incorporate Dataset-JSON Alignment of Metadata with the Define-XML and the e-Protocol as CORE custom rules; Alignment vs BC specializations, e-SAP and Validation of Analysis Results,

Thank You!

Questions?

MHungria@DicoreGroup.com, [AllanH@DicoreGroup.com](mailto>AllanH@DicoreGroup.com), CDISCservices@DicoreGroup.com

www.linkedin.com/in/marcelinahungriadicoregroup

www.linkedin.com/in/marcelinahungria

www.dicoregroup.com

