2025 EXPLORATION FUTURE

Will Dataset-JSON Become the Standard for Next-Generation Data Transport Format?

Jianfeng YE/ SANOFI/ 20250418

Meet the Speaker

Jianfeng YE

Global Submission Solution Leader

SANOFI/SPAR

Disclaimer and Disclosures

- *{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.*

The Answer is

Dataset-JSON

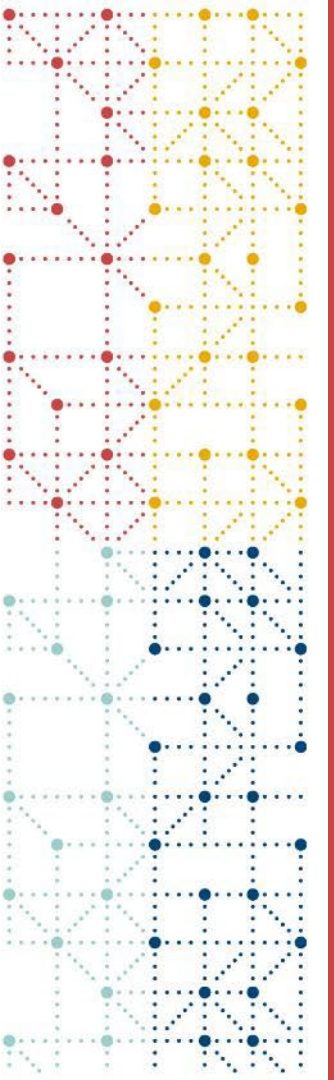
POSSIBLE

有可能

JSON

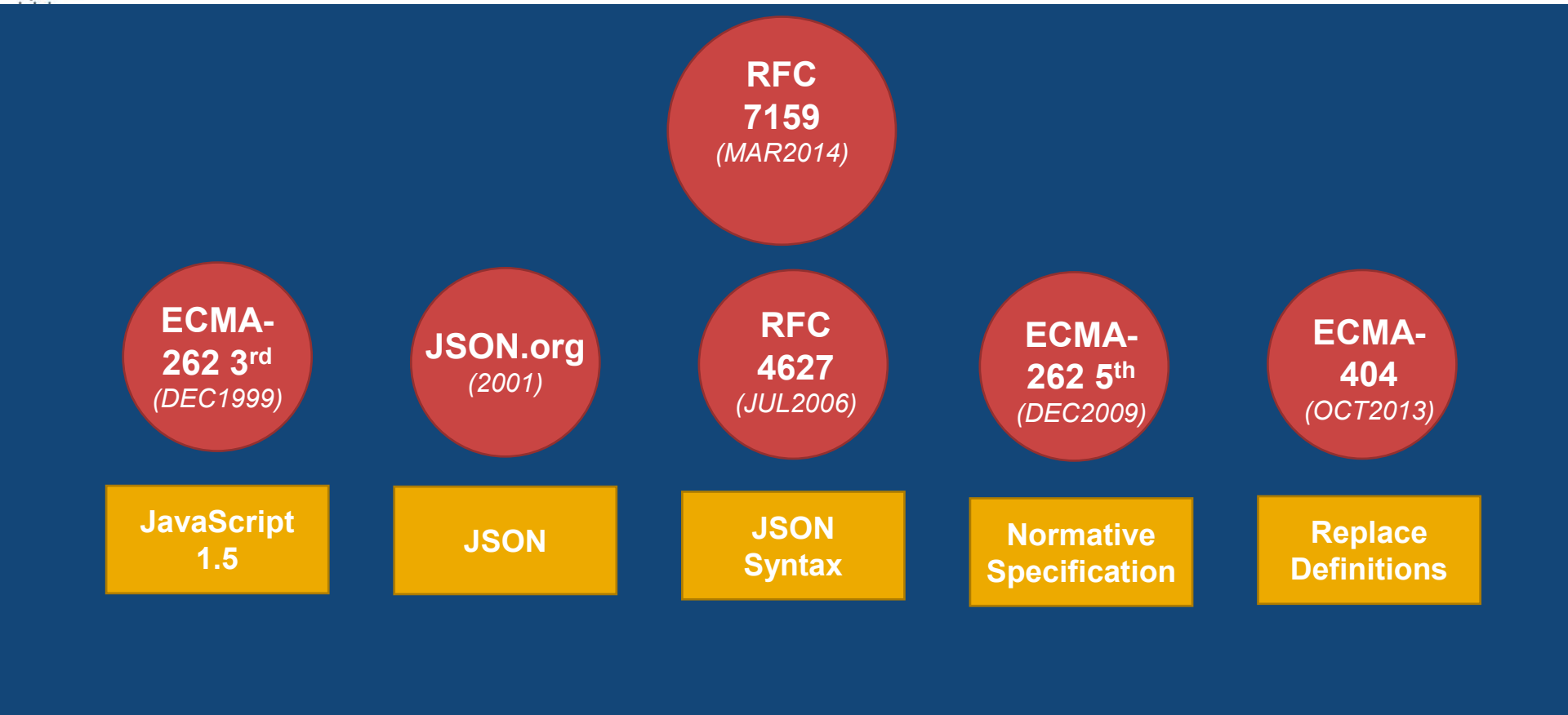
ABSOLUTELY

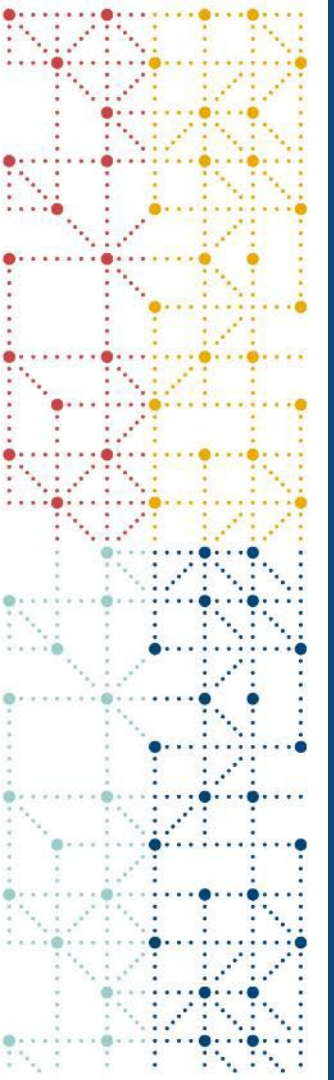
必须的



What is JSON?

JavaScript Object Notation





What are Features of JSON?

ECMA 404： JSON is a text syntax that facilitates structured data interchange between all programming language

JSON is a lightweight, text-based, language-independent syntax for defining data interchange formats. It was derived from the ECMAScript programming language, but is programming language independent. JSON defines a small set of structuring rules for the portable representation of structured data.

STRUCTURAL TOKEN

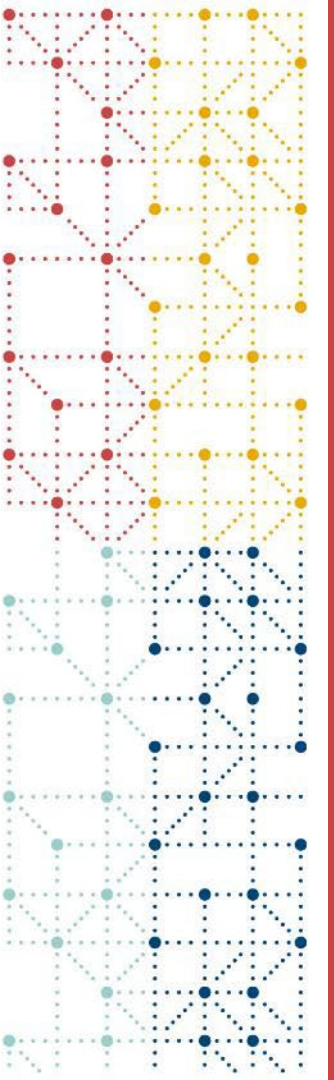
[	U+005B	Left Square Bracket
{	U+007B	Left Curly Bracket
]	U+005D	Right Square Bracket
}	U+007D	Right Curly Bracket
:	U+003A	Colon
,	U+002C	Comma

VALUE

Object	{"name":"Douglas Crockford"}
Array	["Douglas Crockford", "Anna", "Peter"]
Number	68
String	"Douglas Crockford"
Boolean	True* or False*
Null	Null*

* Literal Name Token

```
{
        "name": "Douglas Crockford",
        "age": 68,
        "isWriter": true,
        “publication": [
            "JavaScript: The Good Parts",
            "How JavaScript Works"
        ],
        "address": {
            "street": "Frostbite Fall",
            "city": "Minnesota"
        }
}
```



What is the Conner Store of web Front-end?

The Core Trio of Web Front-end Development

HTML

Hyper Text Markup Language

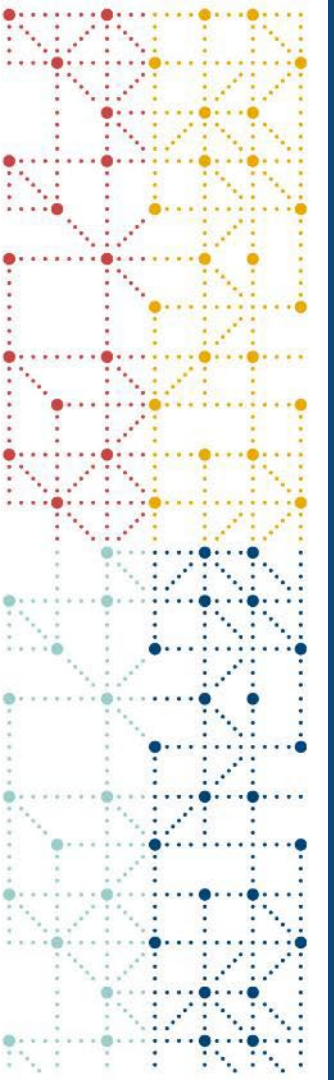
JavaScript

CSS

Cascading Style Sheets

Chrome/Edge/Firefox

Python (Flask) / R (Shiny) / PHP (Laravel) / JAVA (Spring Boot) / C#(.NET)



What is the Dataset-JSON?

Dataset-JSON is a ***data exchange standard*** *for sharing tabular data* ***using JSON****. It is designed to meet a wide range of data exchange scenarios, including* ***regulatory submissions*** *and* ***API-based data exchange****.*

What is the Goal of Dataset-JSON Pilot?

Purpose

PURPOSE

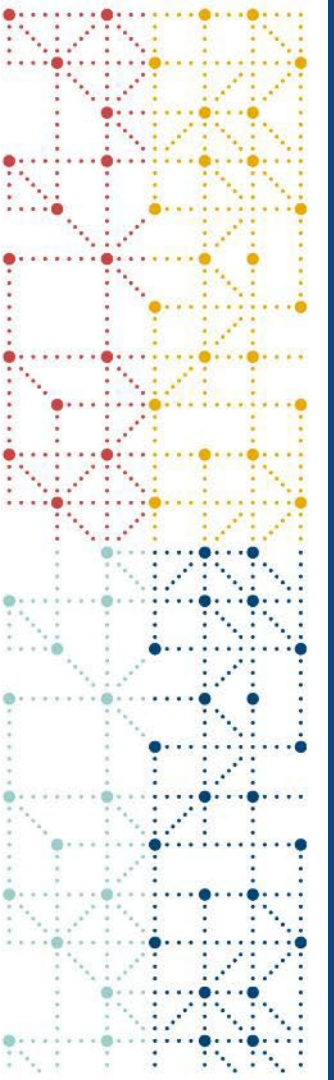
Dataset-JSON as an Alternative Transport Format for Regulatory Submissions Pilot

TARGET

Dataset-JSON 1.0
ODM V2.0

Timeline

MAY2023	SEP2023	OCT2023	DEC2023	APR2024
CDISC/PHUSE Webinar Call for Volunteers.	Cosa Hackathon Kickoff concludes at US interchange.	CDISC US Interchange Plenary Presentation.	Complete Clinical Data Pilot.	Complete Nonclinical Data Pilot.
Pilot Kickoff Meeting.	PHUSE CSS Conference Dataset-JSON Plenary & Workshop.	PHUSE EU Connect Dataset-JSON Workshop.	PHUSE US Connect Pilot Findings Presentation.	PHUSE CSS Final Pilot Report.
JUL2023	SEP2023	NOV2023	FEB2024	JUN2024



What are the Workgroup for this Pilot?

DATA SUBMISSION STANDARDS

BUSINESS CASE

This sub-team will build upon the existing advantages that have been shared for using dataset-JSON to develop a business case for using this format for regulatory data submissions (where they are accepted), as well as developing up an FAQ for sponsors. This will include a call for sponsors to test internally (using CDISC hackathon tools) and to consolidate feedback on these experiences.

TECHNICAL IMPLEMENTATION

XPT is a long established de-facto standard and changing it is likely to have many knock-on consequences. This sub-team will explore the practical implications of migrating to Dataset-JSON, and develop a technical roadmap, guidance and support materials for industry

Strategy for Future Development

This sub-team will discuss and make proposals of how to take advantage of the enhancements dataset-JSON brings over xpt as a data transfer format. It will consider how to balance continuity vs a more forward-looking approach to the design of data standards and the submission process

PILOT SUBMISSIONS REPORT

This sub-team will consolidate feedback from the pilot submissions and develop a report on the feedback received from those participating in the pilot submissions.

What are the Stakeholders for this Pilot?

REGULATORY

FDA
EMA
PMDA
NMPA
Health Canada
DCGI

PHARMA

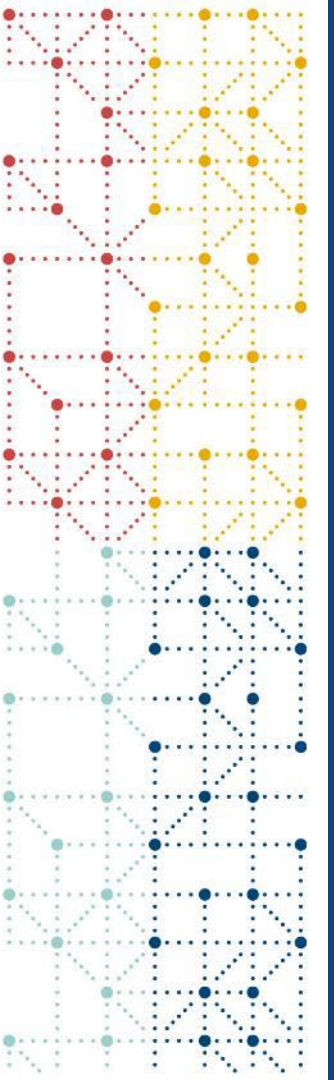
Alnylam Pharmaceuticals
BioCryst
BMS
GSK
J&J
Merck
Novo Nordisk
OTSUK
Roche
Sanofi
Sumitomo Pharma America
Takeda
Vertex

CRO

AGATI clinical
AQB
ICON
IQVIA
PAREXEL

TOOL VENDER

Certara
CORE from CDISC
COSA repositor
Domino
eClinical
Edetek
Medidata
Nonmem software for clinpharm trials
Pinnacle 21/Formedix
Saama
SAS/LSAF



What is the Pilot Strategy?

INITIAL TESTING

Industry and FDA complete internal testing utilizing CDISC hackathon tools

EXTERNAL TESTING

Test regulatory JSON submission via test electronic study gateway

EVALUATE FINDINGS

The Team will review findings from the questionnaire and FDA testing

CONCLUSION

Report out on findings to industry and address issues in Dataset-JSON v1.1

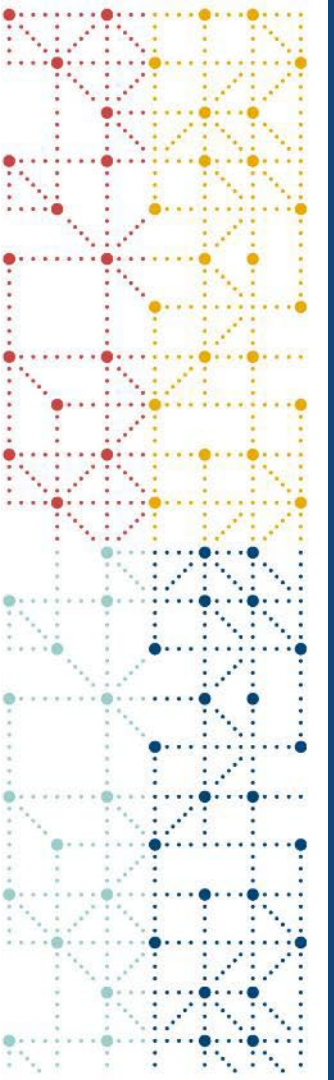
What are the Pilot Results?

SUMMARY

Overall results minor data representation, display format issue, and precision concerns.

ACTION

- **Standards updates, user guide content, tool update and enhancements.**
- **Many findings related to the conversion tools and interoperability testing.**
- **Most issues related to conversion tools and interoperability testing across different tools.**

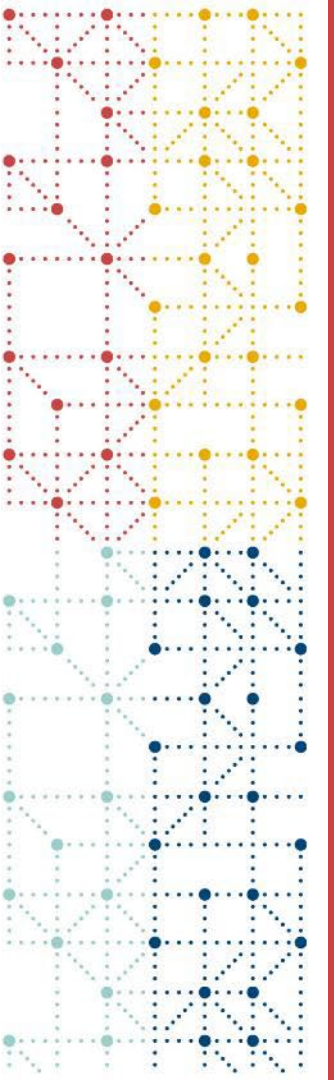


What are the Pilot Conclusion?

Dataset-JSON can transport information with no disruption to business and is viable as the primary transport option.

Dataset-JSON files successfully tested and submitted to FDA via test electronic Study Gateway with no data integrity issues

Dataset-JSON v1.1 to address the pilot finding



What are the

Next

after this Pilot?

Complete the Pilot Report

Define Dataset-JSON v1.1 and user guide

Update the convention and viewer software

Publish the Dataset-JSON API standard

Collaborate with FDA /PMDA to define the next steps for regulatory support

What are Features of Dataset-JSON v1.1?

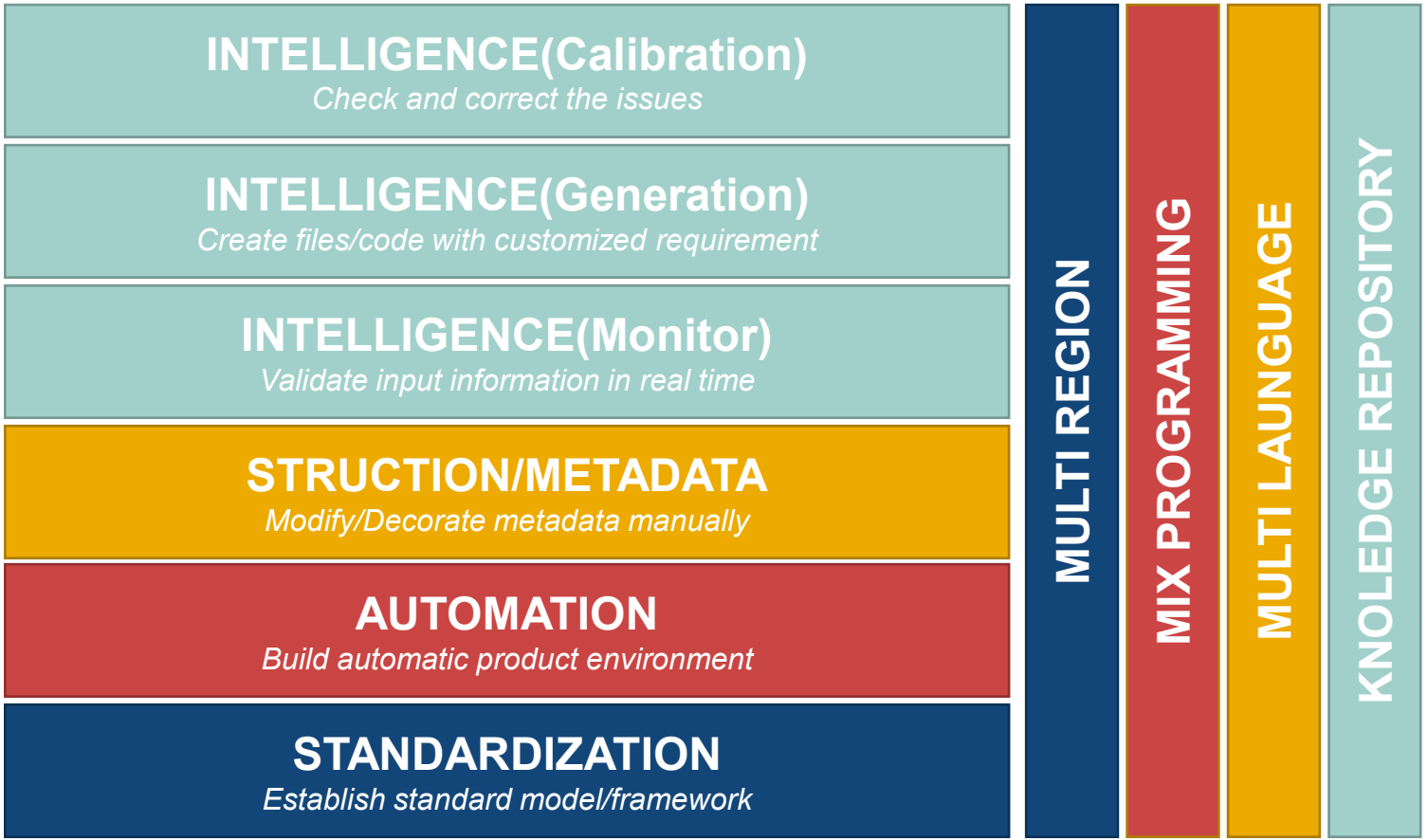
Each Dataset-JSON dataset can optionally reference a Define-XML document containing more complete metadata for the dataset.

Dataset-JSON is based on the JSON standard that is ***simple*** *to implement, very* ***stable****, and* ***widely supported****.*

It is also designed to address the limitations of legacy formats and is extensible to support new metadata and new use cases.

Retrospect to 2023/8/25

CDISC China Interchange


INTELLIGENCE(Calibration)
Check and correct the issues
INTELLIGENCE(Generation)
Create files/code with customized requirement
INTELLIGENCE(Monitor)
Validate input information in real time
STRUCTION/METADATA
Modify/Decorate metadata manually
AUTOMATION
Build automatic product environment
STANDARDIZATION
Establish standard model/framework
MULTI REGION
MIX PROGRAMMING
MULTI LAUNGUAGE
KNOLEDGE REPOSITORY

LOW-CODE
HIGH-INTERACTION
AI ASSISTANCE

NETWOK FRIENDLY DATA/FILE	FORMATING TLF	ASSISTANT APPROVAL/TRIAL DESIGN
Storage Exchange Visualization *(DATASET-JSON/XML)*	Easy Review Customized Result Diversify Exhibition Faster Approval	Data Issue Indicator Abnormal Potential Risk Prevention Cross Check *(STUDY/PHASE/TA)*

REFERENCE

FDA https://www.fda.gov/
<Study Data Standards Resources>
<Dataset-JSON Pilot Report and Next Steps>

CDISC https://www.cdisc.org/
<Dataset-JSON v1.1>
<Dataset-JSON v1.0 >

FEDERAL REGISTER https://www.federalregister.gov/
Dataset-JSON Pilot Report and Next Steps Jesse Anderson, FDA-CDER Sam Hume, CDISC

ECMA https://ecma-international.org/
<ECMA 262>
<ECMA 404>

RFC https://www.rfc-editor.org/
<RFC 4627>
<RFC 7159>

JSON.org https://www.json.org

Thank You!

cdisc