

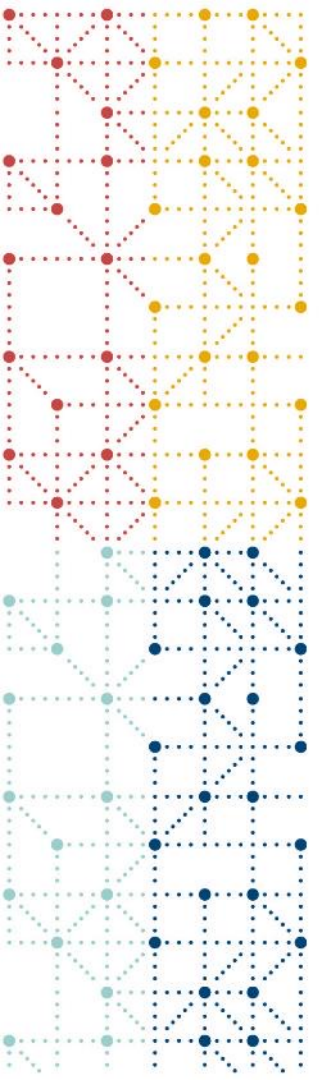


# COSA Dataset-JSON Hackathon II Results

Sam Hume

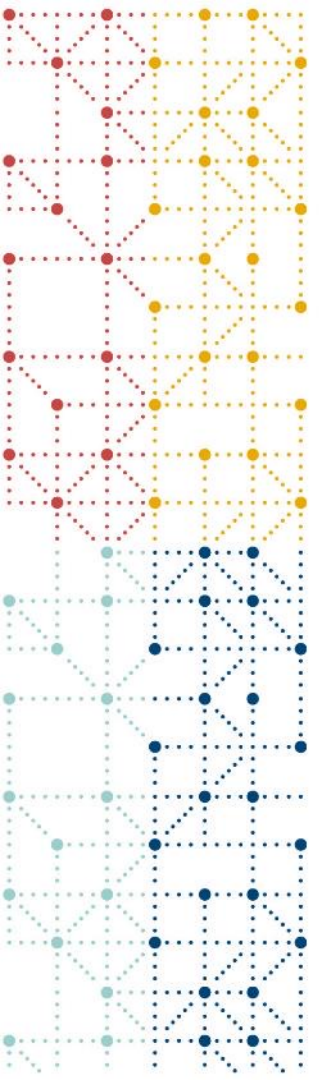
2023-03-26





## Agenda

1. Hackathon objectives and background
2. API specification development
3. Next steps



# Hackathon Objectives and Background



# Problem Statement

- Primary objective: Create a draft REST API specification for Dataset-JSON
- Secondary objective: Proof-of-concept implementations to demonstrate and test the API specification
- Virtual hackathon
  - The team worked collaboratively to develop and test the draft specification
  - Approximately 100 people registered to participate
  - Results communicated at the 2023 US Interchange
- Dates: Sept. 1 – October 13, 2023

# What do we plan to do with the draft API specification?

The draft API specification is available for review and comments

The API draft specification will be delivered to the Dataset-JSON Team for eventual review, refinement, and publication

API Spec will be communicated to the Dataset-JSON Pilot teams

Implement prototypes to demonstrate and test the API specification

# What are the goals of the pilot?

## Milestone 1: Short Term

- Pilot using JSON format with existing XPT ingress/egress to carry the same data
- Same content, different suitcase, no disruption to business process on either side
- Allow FDA to evaluate how internal tools can support JSON format

➔ **Success Criteria: Demonstrate that Dataset-JSON can transport information with no disruption to business**

## Milestone 2: Development of future strategy

- Evaluate how current and future industry standards can benefit without XPT limitations  
e.g., Variable names > 8, labels > 40, data > 200
- Evaluate combining metadata with data  
e.g., Define-XML / Define-JSON based
- Enhanced conformance rules
- FDA to utilize findings to evaluate tool redevelopment plan to natively consume files in JSON format

➔ **Success Criteria: Demonstrate the viability of Dataset-JSON as the primary transport option**



# Collaborative Work

- GitHub repository created to develop the API spec collaboratively
  - <https://github.com/cdisc-org/DataExchange-DatasetJson-API>
  - OpenAPI 3.1 is used as the standard for the machine-readable spec
  - Project communication took place in Slack and GitHub discussions
- An OpenAPI Specification schema is available to aid in development
  - <https://github.com/OAI/OpenAPI-Specification>
- Numerous tools, including many free ones, work with OpenAPI
  - <https://openapi.tools/>
  - Code generators, editors, data validators, and many other tools
- Demonstration API implementations will be available



# API Specification Development

Dataset-JSON API draft specification development during the hackathon





# API Scope & Assumptions

---

Focused on data exchange

---

Support the basics with the intention that this becomes a standard

---

Based on the same version of Dataset-JSON used in the pilot

---

May use compression and/or pagination to support larger datasets

---

Supports a broad range of data exchange scenarios beyond submissions

# API draft documentation generated from specification

GET	/about	Get About the API
GET	/studies	Get the list of studies
POST	/studies	Create a new study
GET	/studies/{studyOID}	Get a study
PUT	/studies/{studyOID}	Update a study
DELETE	/studies/{studyOID}	Delete a study
GET	/studies/{studyOID}/datasets	Get the list of datasets for a study
POST	/studies/{studyOID}/datasets	Create a new dataset
GET	/studies/{studyOID}/datasets/{datasetOID}	Get a dataset
PUT	/studies/{studyOID}/datasets/{datasetOID}	Update a dataset
PATCH	/studies/{studyOID}/datasets/{datasetOID}	Append records to a dataset
DELETE	/studies/{studyOID}/datasets/{datasetOID}	Delete a dataset

# API draft documentation details

## Request

**GET** `/studies/{studyOID}/datasets` Get the list of datasets for a study

**Parameters**

Name	Description
<b>studyOID</b> <small>* required</small>	<input type="text" value="studyOID"/>
<b>string</b> <small>(path)</small>	
<b>standard</b> <small>(query)</small>	<input type="text" value="standard"/>
<b>if-modified-since</b> <small>(header)</small>	<small>Default value :</small> <input type="text" value="if-modified-since"/>

## Response

Code	Description
200	List of datasets for a study
	Media type <input type="text" value="application/json"/>
	<small>Controls Accept header.</small>
	<a href="#">Example Value</a>   <a href="#">Schema</a>
	<pre>[   {     "oid": "string",     "name": "string",     "label": "string",     "type": "string",     "records": 0,     "href": "string",     "filename": "string"   } ]</pre>

# Dataset Compression and Pagination

**GET** /studies/{studyOID}/datasets/{datasetOID} Get a dataset

**Parameters**

Name	Description
<b>studyOID</b> <i>* required</i>	<input type="text" value="studyOID"/>
string (path)	
<b>datasetOID</b> <i>* required</i>	<input type="text" value="datasetOID"/>
string (path)	
metadataonly (query)	Default value : false
	<input type="text" value="metadataonly"/>

offset  
(query) Default value : 0

limit  
(query) Default value : 0

accept-encoding  
(header) Default value :

if-modified-since  
(header) Default value :

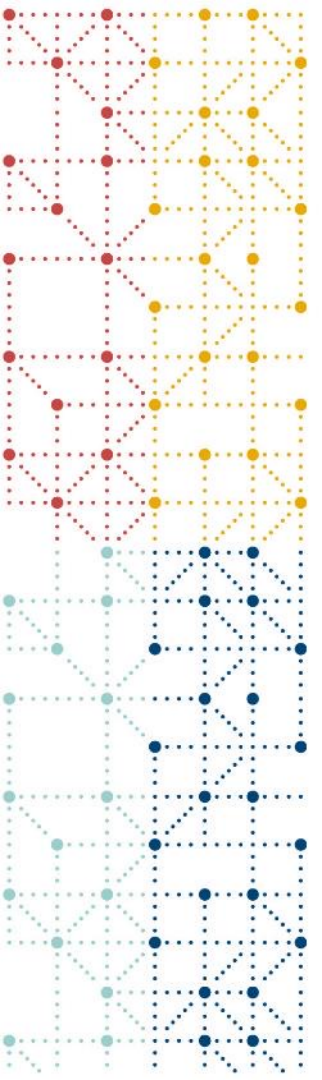


## Next Steps



# Next Steps

- Complete the PHUSE/CDISC/FDA Dataset-JSON as an alternative transport format for regulatory submissions pilot
  - Read out of the final report scheduled for the PHUSE CSS conference June 3-5
- Make Dataset-JSON API prototype available to try
  - Planned for Q2 2024
- Develop Dataset-JSON v1.1 as a standalone standard
  - Updates include the findings based on the pilot
  - Plan to schedule a public review of Dataset-JSON v1.1 for Q4 2024
  - Call for volunteers was distributed last week
- Plan to publish the Dataset-JSON API shortly after publishing Dataset-JSON v1.1
  - The API is essential to data exchange using Dataset-JSON
  - Comments on the draft API specification welcome



## Additional Resources

# Resources

- Dataset-JSON API draft specification
  - <https://github.com/cdisc-org/DataExchange-DatasetJson-API>
- Dataset-JSON specification
  - <https://www.cdisc.org/dataset-json>
  - <https://wiki.cdisc.org/display/PUB/Dataset-JSON>
- Dataset-JSON GitHub repository
  - <https://github.com/cdisc-org/DataExchange-DatasetJson>
- Dataset-JSON v1.1 wiki
  - <https://wiki.cdisc.org/display/DSJSON1DOT1/Dataset-JSON+v1.1>
- COSA Directory Dataset-JSON Hackathon I projects
  - <https://cosa.cdisc.org/hackathons/datasetJson>



# Thank You!

## Questions?

Interested in volunteering for Dataset-JSON v1.1?

<https://www.cdisc.org/volunteer/form>

[shume@cdisc.org](mailto:shume@cdisc.org)

<https://www.linkedin.com/in/sam-hume-dsc>

