



CSS 2024

The Computational Science Symposium

3-5 June
Silver Spring, MD



phuse.global

Dataset-JSON as Alternative Transport Format for Regulatory Submissions

Jesse Anderson (FDA/CDER)
Jesse.Anderson@fda.hhs.gov

Stuart Malcolm (PHUSE/Veramed)
stuart.Malcolm@veramed.co.uk

Sam Hume (CDISC)
shume@cdisc.org

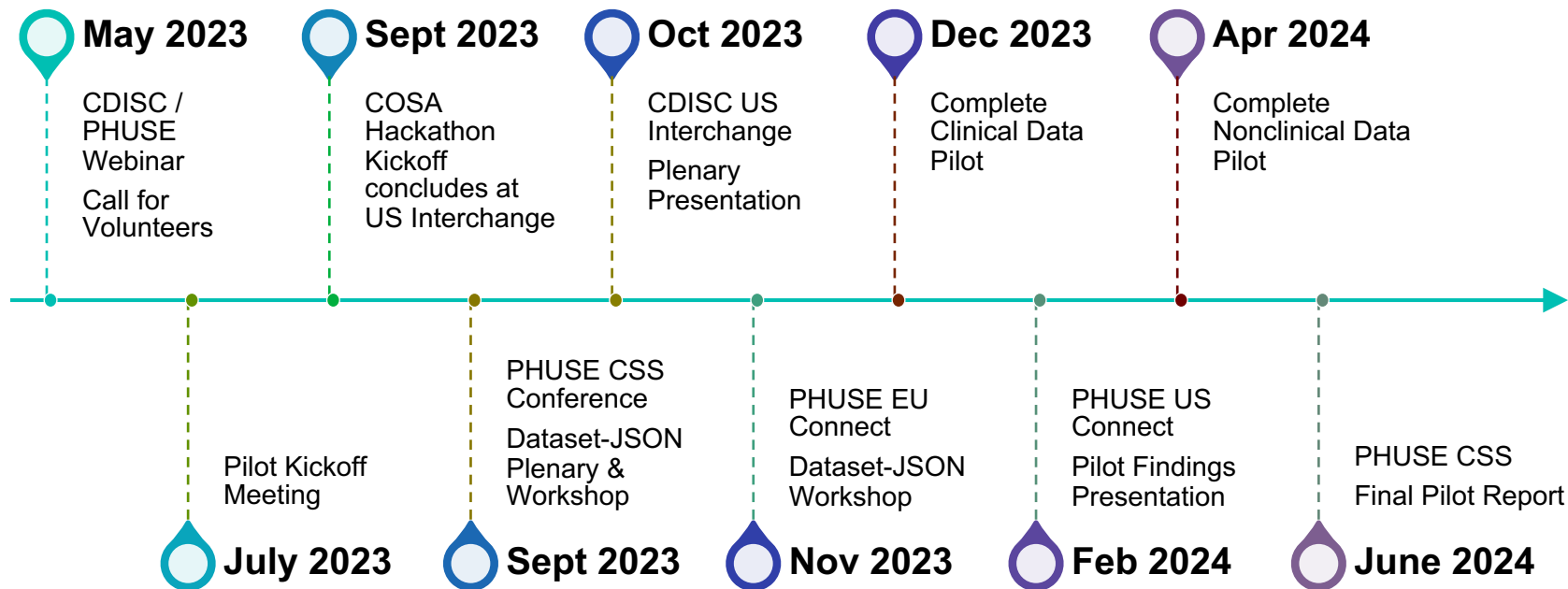


Introduction

- Collaboration between PHUSE, CDISC and FDA
- Build upon the considerable amount of work done over the years to replace XPT as the default file format for clinical and device data submissions to regulatory authorities.
- Four Sub-teams
 - Pilot Submissions Report
 - Technical Implementation
 - Business Case
 - Strategy for Future Development



Project Timeline





What is Dataset-JSON?

- CDISC created dataset exchange standard for exchanging tabular data leveraging JSON format and designed to meet regulatory submission needs
- Advantages:
 - Open-source and truly human readable
 - Schema supports any tabular format
 - Eliminates limitations of legacy formats
 - Extensible to support new metadata and new use cases



What were 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**



FDA CDER OCS Pilot Strategy

Initial Testing

External Testing

Evaluate Findings



Initial Testing: Industry and FDA complete internal testing utilizing CDISC Hackathon Tools



External Testing: Test regulatory JSON submissions via test Electronic Study Gateway



Evaluate Findings: Team to review findings from questionnaire and FDA testing



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



Summary of Findings



Overall, results showed minor date representation, display format issues, and precision concerns (full findings listing [here](#))



Findings can be addressed with: (1) standards updates, (2) User Guide content, and (3) tool updates and enhancements



Many findings related to the conversion tools and interoperability testing



Most issues (e.g., date representation, precision) related to conversion tools and interoperability testing across different tools



Pilot Conclusions



Milestone 1 satisfied: Dataset-JSON can transport information with no disruption to business and is viable as the primary transport option



FDA testing noted that, to date, COTS analytical tools do not accept JSON files. Updates necessary for full implementation of Dataset-JSON as transport file for regulatory data.



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 findings within the next year



Dataset-JSON v1.1 Addresses the Pilot Findings

Update

Update the Dataset-JSON Standard

- E.g., add NDJSON support and targetDataType
- Update the specification, schema, and examples

Create

Create a User Guide

- Many findings requests for information
- The User Guide will be a living document

Enhance

Enhance and update software tools

- Conversion software
- Viewer software



Open-Source Conversion Software Tools



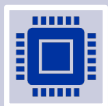
SAS

- [The SAS conversion software by Lex Jansen](#)



R

- [R conversion package by Atorus Research and Johnson & Johnson](#)



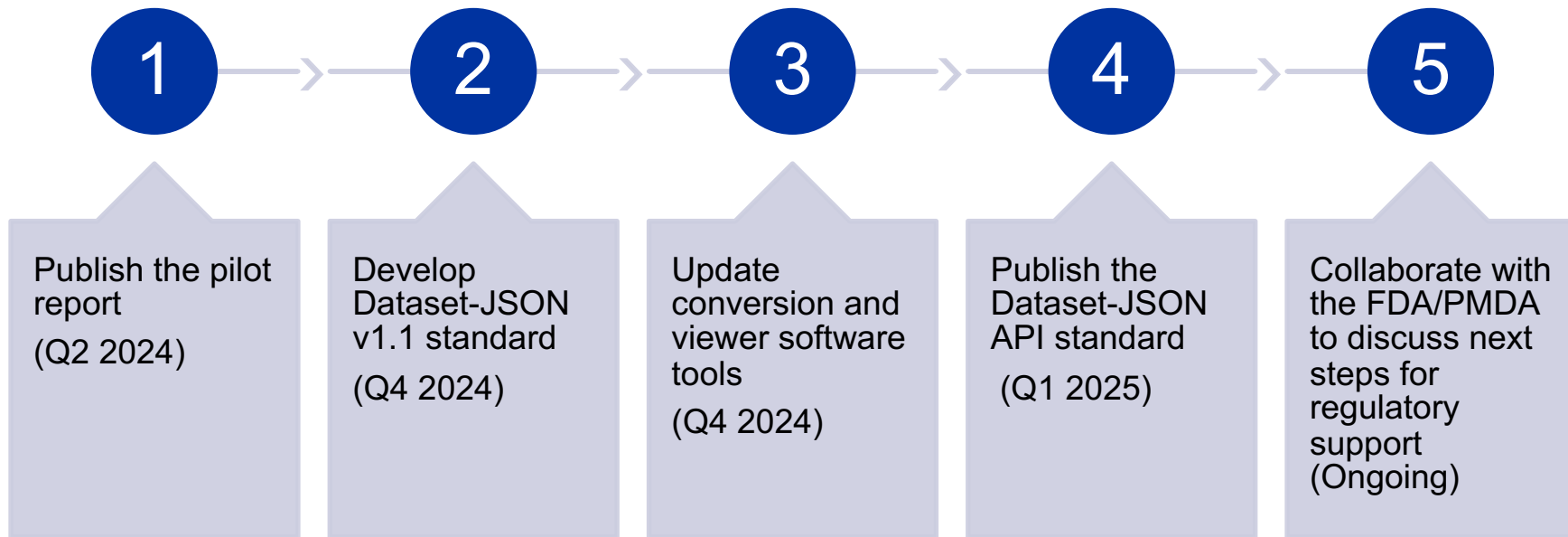
Python

- Multiple Python conversion software tools

- Volunteers committed to ensuring a Parquet conversion tool is available
- Open-source software teams need contributors



Next Steps



Thank You!
Questions?

