

Dataset-JSON as Alternative Transport Format for Regulatory Submissions

Sam Hume (CDISC) and Stuart Malcom (PHUSE/Veramed)
shume@cdisc.org stuart.Malcolm@veramed.co.uk

Pilot Overview & Milestones

CDISC, PHUSE, and the FDA proceeded with a pilot to demonstrate that Dataset-JSON can deliver an open-source, human readable, and flexible transfer model to eliminate XPT limitations. The milestones for the project were as follows:

Milestone 1 (Short Term)

- JSON pilot submissions with existing XPT ingress/egress
- Evaluate same content, different suitcase
- **Success Criteria:** Demonstrate Dataset-JSON transports information with no disruption to business

Milestone 2 (Long Term)

- Evaluate how current & future standards can benefit without XPT limitations
- Evaluate combining metadata with data files
- **Success Criteria:** Demonstrate viability of Dataset-JSON as primary transport option

Methods

Pilot participants were asked to perform the below steps utilizing existing software conversion tools (e.g., SAS, R, Python) and report their results via a questionnaire.

1. Select the software conversion tool(s) to use (one or more)
2. Select study datasets and convert them to Dataset-JSON
3. Convert Dataset-JSON datasets into a native dataset form
4. Compare the native datasets back to the original format and ensure there are no unexpected changes in the converted dataset
5. Optionally, test other conversion scenarios (e.g., interoperability tests)
6. Complete the on-line questionnaire to record your findings

Results

Overall, Dataset-JSON can function as an XPT alternative

Dataset-JSON met the pilot objectives:

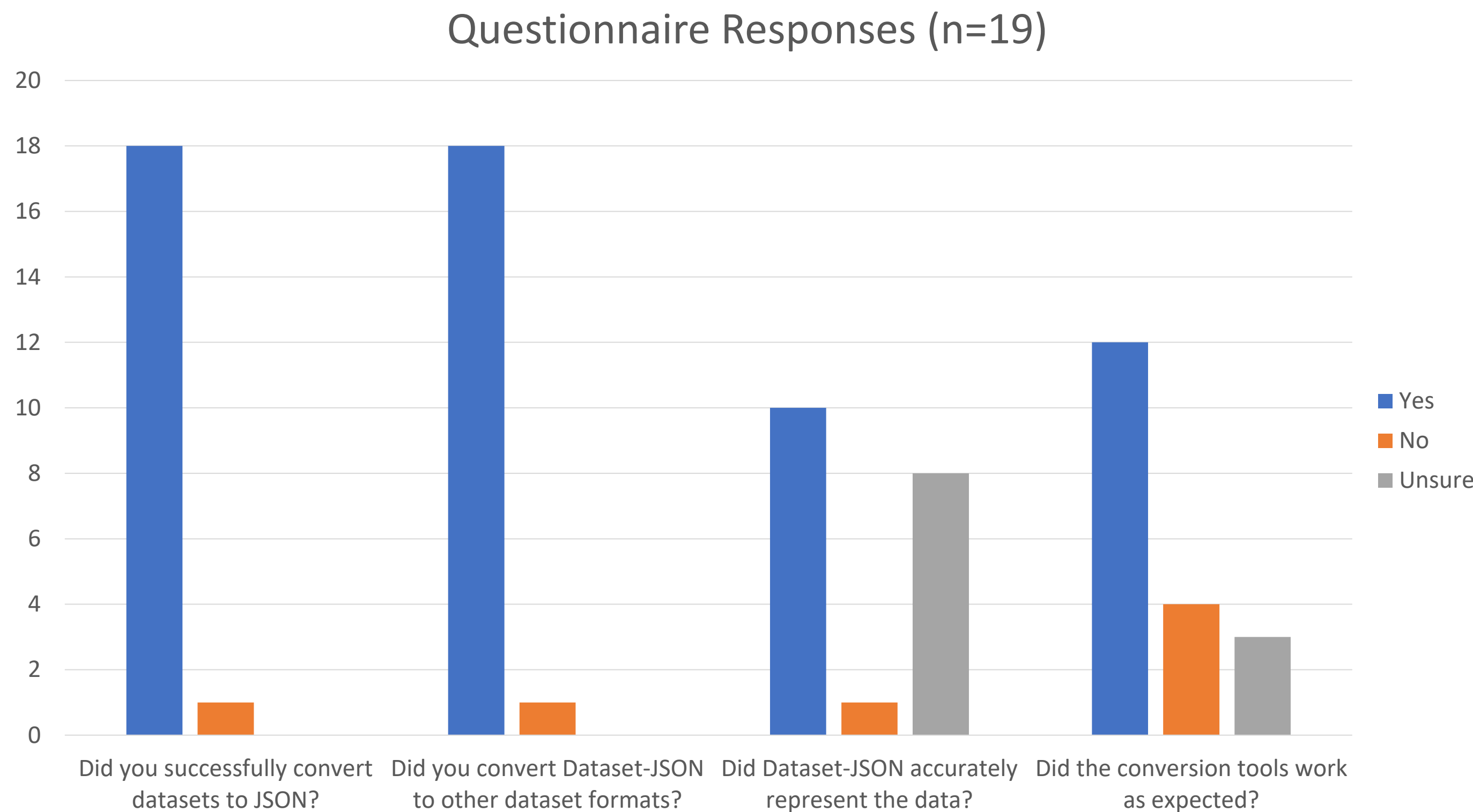
- No show-stoppers were identified
- Milestone 1 was satisfied
- FDA testing was successful

Dataset-JSON v1.1 implements improvements based on the findings

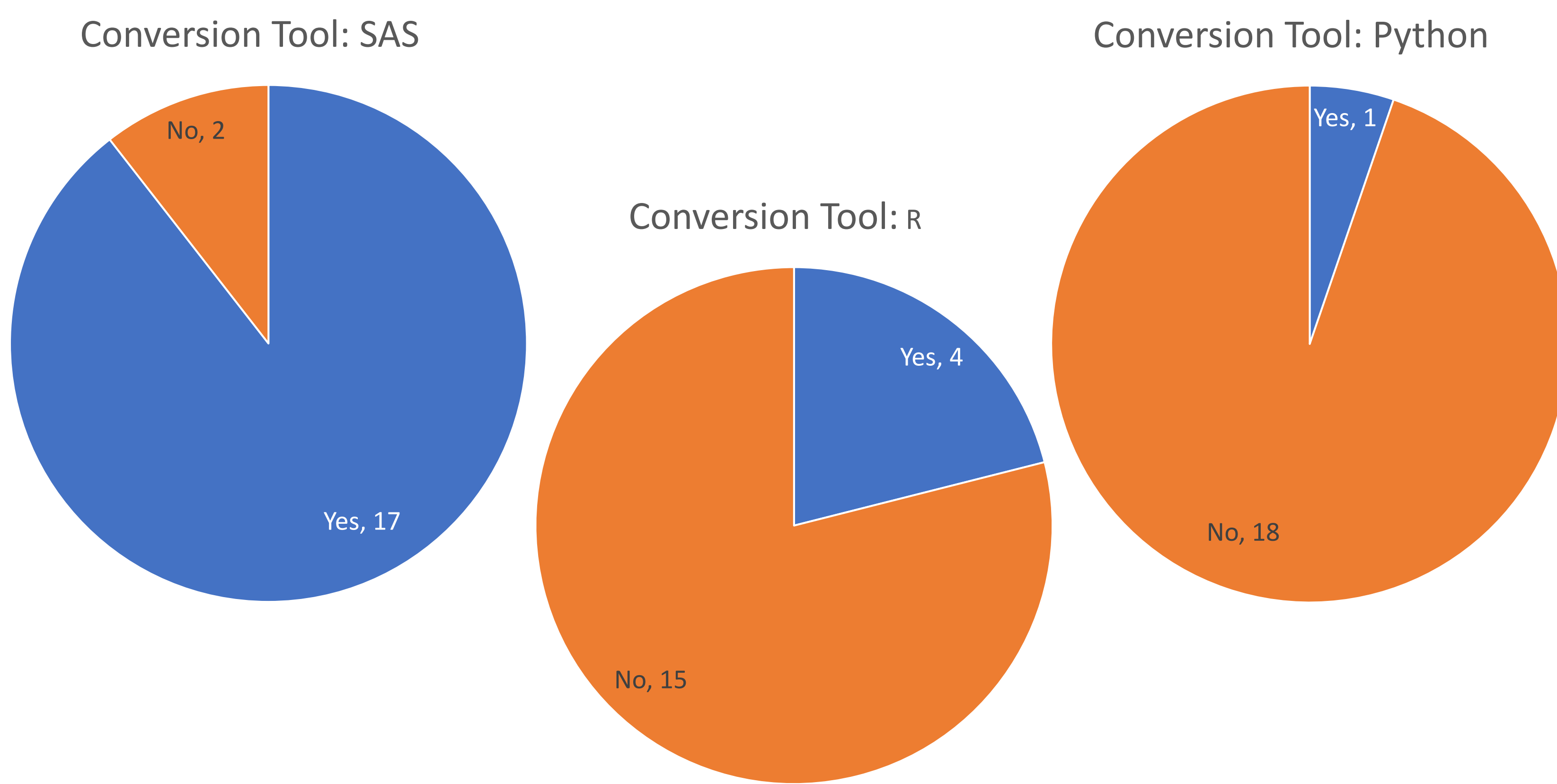
Three Categories:

1. Update the Dataset-JSON standard (e.g., add NDJSON support)
2. Create a User Guide
3. Update the conversion tools (e.g., better performance processing large datasets)

Questionnaire Responses

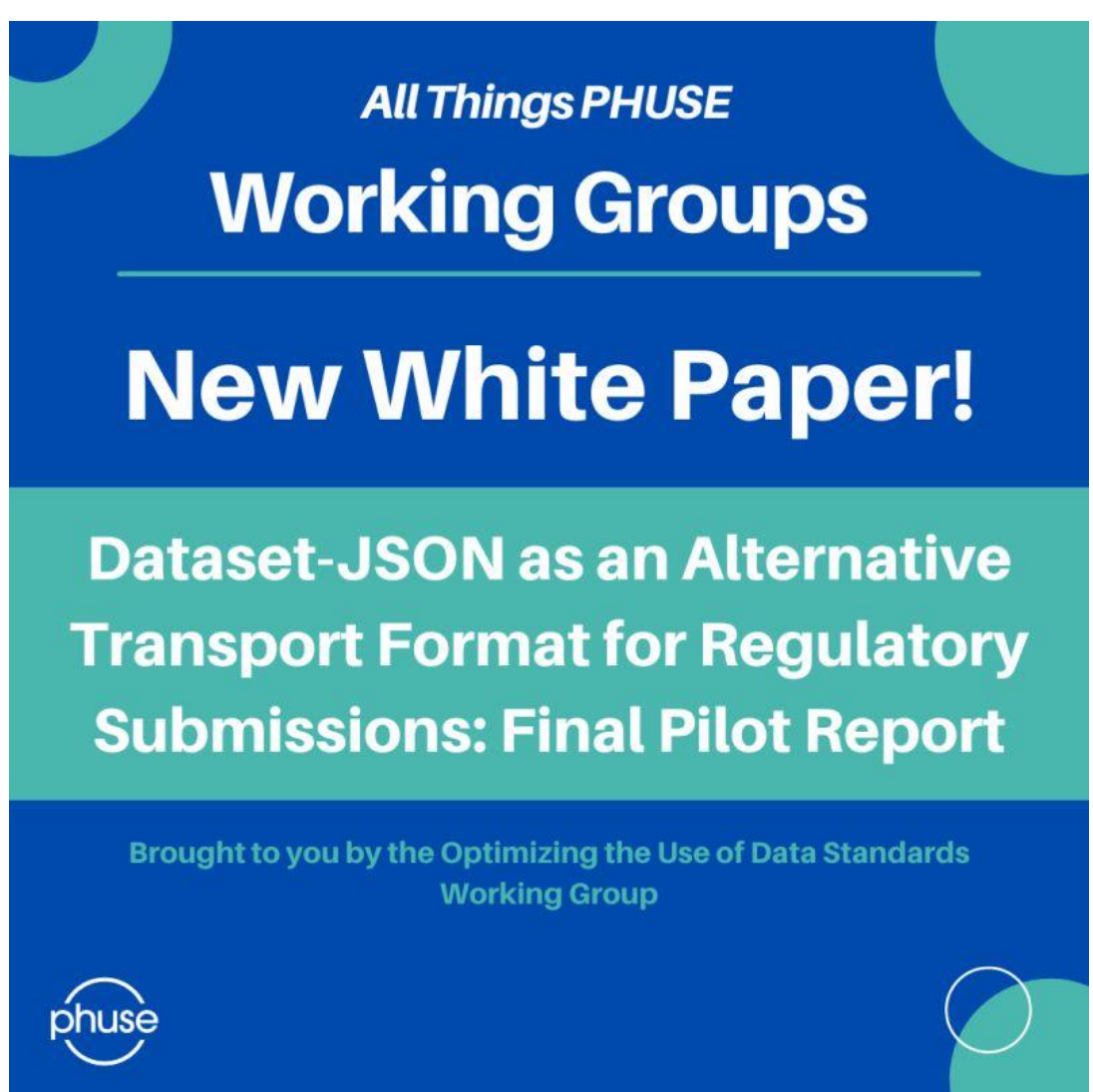


Tools Used:



Final Pilot Report

- The pilot findings support using Dataset-JSON as an alternative transport format for XPT v5
- Solutions to the findings will be implemented in Dataset-JSON v1.1
- Published 19 June 2024
- Doc ID: WP 88



<https://phuse.s3.eu-central-1.amazonaws.com/Deliverables/Optimizing+the+Use+of+Data+Standards/WP-88+Dataset-JSON+Report.pdf>

Next Steps

Complete Dataset-JSON v1.1 Public Review (Oct. 2024)

Publish the Dataset-JSON API standard (Q2 2025)

