Dataset-JSON Viewer Hackathon Kick-off

Sam Hume, CDISC

Charles Shadle, CDISC

2024-10-08

Agenda

1. Introduction
2. Objectives
3. Judging Criteria and Rules
4. Timeline
5. Resources
6. Communications
7. Questions

Introduction

Introducing the COSA Dataset-JSON Viewer Hackathon

Welcome to the COSA Dataset-JSON Viewer Hackathon

~ 200 registered participants	~ 12 weeks in duration	Report outcomes Q1 COSA Spotlight
Results presentation at EU Interchange	Viewers are a *Must Have* requirement	A competitive hackathon

cdisc®

CDISC Open-Source Alliance (COSA)

COSA Mission: The CDISC Open-Source Alliance (COSA) supports, promotes, and sometimes sponsors open-source and free software development projects that create tools for implementing or developing CDISC standards to drive innovation in the CDISC community.

- COSA has hosted several hackathons
  - This is the 3rd Dataset-JSON Hackathon
- Dataset-JSON Hackathon solutions may apply to be included in the COSA Repository Directory
  - Projects from the first Dataset-JSON Hackathon: https://cosa.cdisc.org/hackathons/datasetJson
- Requires an open-source license
- Requires a public repository
- There will be opportunities to demonstrate your viewer software
 - COSA Spotlight Webinar to demo solutions

https://cosa.cdisc.org/

Objectives

The Hackathon Objectives

COSA Dataset-JSON Viewer Hackathon

- <u>Primary objective</u>: Create Dataset-JSON Viewer software
  - Projects can be new or extensions to existing software
  - Projects must support Dataset-JSON v1.1

- Significant and broad interest in viewer software for Dataset-JSON
  - A finding from the *Dataset-JSON as an Alternative Transport Format for Regulatory Submissions Pilot*

- Virtual hackathon
  - Individuals or teams will work virtually
  - Periodic meetings will be scheduled to discuss progress and answer questions
  - Project demonstrations will be scheduled for a COSA Spotlight Webinar

- Dates: Oct. 8th – Jan. 6th
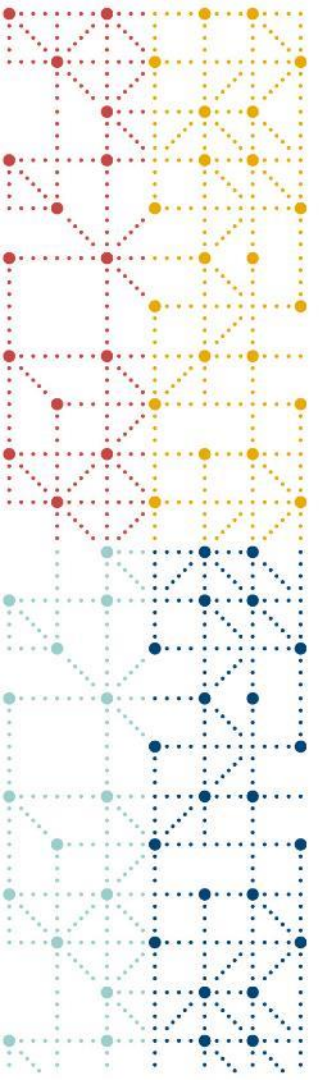
Judging Criteria and Rules

The Hackathon judging criteria and rules that each project should follow and how each project will be evaluated by the judges

Hackathon Judging Criteria

- Usability
  - Overall ease of use
  - Ease of getting started

- Functionality
  - Key features
    - Basic browsing, filtering, and sorting
  - Extended features

- Deployment
  - Ease of installation
  - Deployment flexibility
  - Ease of access

- Performance
  - Responsiveness
  - Ability to handle large datasets

- Documentation
  - Documentation of viewer features
  - Instructions on using the viewer
  - Changes from baseline

- Bonus
  - NDJSON support
  - Support for the final version of Dataset-JSON v1.1 (include changes made after Public Review)
  - Design (appearance)

cdisc.

Hackathon Rules and Guidelines

- The project must be available in a public GitHub repository (or equivalent)
- All judged entries must be released under an open-source license
- We must be able to run the viewer in order to judge it
  - If the project is cloud-based, please provide access to a hosted version for judging
  - If the project is locally deployed, please provide installation instructions
- Provide a README.md that highlights the viewer's features and provides instructions for using the Viewer
- At a minimum, it must support the Public Review version of Dataset-JSON v1.1
- Support for Dataset-JSON v1.0 is not expected and will not be evaluated.
- If updating an existing Viewer, please report what was added or improved as part of the hackathon

cdisc.

Competitive Hackathon

- 3 winners will be selected

- Winners will receive prizes
    - Swag (e.g. t-shirt)
    - Certificate
    - Bragging rights
    - Q1 COSA Spotlight recognition and demonstration + additional promotions and demos
    - Possibly some additional sweeteners

- The COSA Governance Board will serve as judges
    - Board members that are participating in the hackathon may not be a judge

- All participants will get
    - Feedback from the judges
    - Basic swag (e.g. sticker)
    - Opportunity to demo their product and become part of COSA

cdisc.

Timeline

Dataset-JSON Viewer Hackathon Timeline

Oct 1-11 2024

Hackathon registration

Jan 6, 2025

Conclude the Dataset-JSON Viewer Hackathon

Mar 2025

Q1 COSA Spotlight 2025 winner recognition and demo

Hackathon Kickoff

Oct 8, 2024

Judges review the entries

Jan 7-10, 2025

EU Interchange winner recognition

April 23-24, 2025

cdisc
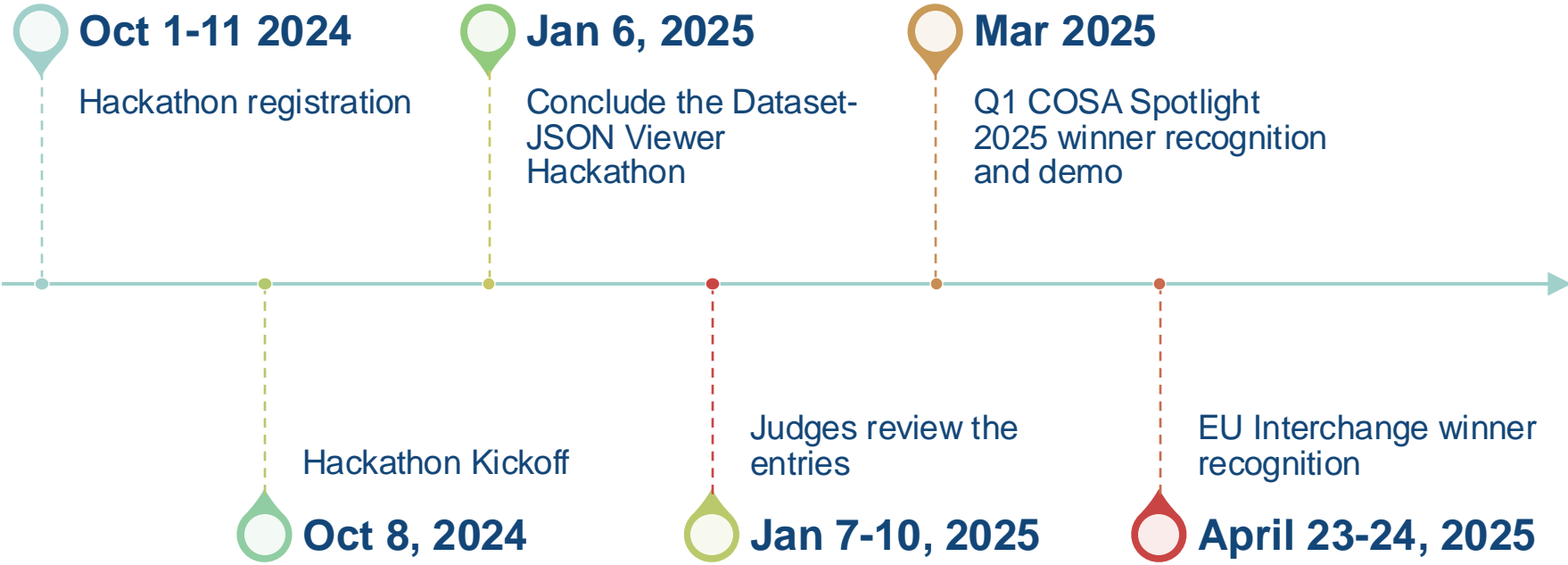
Public Review

- Public Review starts on September 10, 2024

- Public Review lasts 30 days and closes on October 10, 2024
  - Some parties have asked for extra time

- The JIRA project is Dataset-JSON review comments (DSJSONCT) located at:
  - https://jira.cdisc.org/projects/DSJSONCT

- As of today, no major changes expected from the Public Review comments
  - This could change based on new comments coming in

- If minor changes are implemented, ideally Hackathon projects will update their viewers to support the final version of Dataset-JSON v1.1
  - If there are changes in Dataset-JSON v1.1 that you do not implement, make sure you make test datasets available and let the judges know when you plan to update the project to support the final v1.1 release

cdisc.

Resources

Links to resources that may be helpful for your Hackathon project

Dataset-JSON v1.1 Public Review Content

1. Dataset-JSON v1.1 Specification
   - https://wiki.cdisc.org/display/DSJSON1DOT1/Dataset-JSON+1.1

2. User Guide
   - https://wiki.cdisc.org/display/DSJSON1DOT1/User%27s+Guide
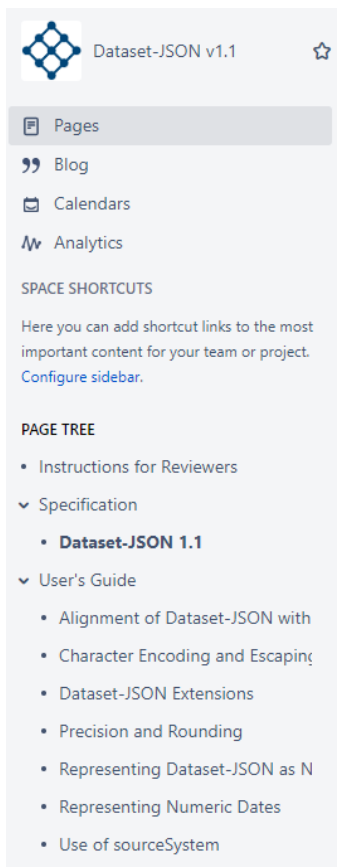
3. Schemas and Example Datasets
   - https://github.com/cdisc-org/DataExchange-DatasetJson

cdisc.

CDISC Wiki

Contains the Dataset-JSON v1.1 Specification and a User's Guide with articles

GitHub: JSON Schema and Examples

- The JSON schema, both for JSON and NDJSON representations, and examples can be found at the GitHub repository for the Dataset-JSON Version 1.1 standard: https://github.com/cdisc-org/DataExchange-DatasetJson
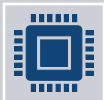
Software Conversion Tools

SAS
- [The SAS conversion software by Lex Jansen](#)
- Includes a macro for comparing libraries with SAS datasets
- Documentation is included

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

Python
- Multiple Python conversion software tools
- Documentation is included
- Covers multiple dataset formats, including Parquet and SAS

COSA Directory Dataset-JSON Hackathon I projects
https://cosa.cdisc.org/hackathons/datasetJson

Communications

Tools for communications during the Hackathon

Hackathon Communication Channels

- To receive communications and meeting notices, please register:
  - https://www.cdisc.org/events/webinar/dataset-json-viewer-hackathon

- Meeting date, time, cadence:
  - Meet-ups every other week (attendance is optional)
  - Date/time of meetings?

- Use existing Dataset-JSON Hackathon Slack workspace
  - Slack is the best way to have an ongoing dialog with the Dataset-JSON Hackathon community
  - Will send out invites to participants not already on it
  - https://join.slack.com/t/dataset-jsonhackathon/shared_invite/zt-2rzvlout0-~e4u2QjTJZ~XNgqSla0eFA

- Wiki site for Dataset-JSON Hackathons
  - https://wiki.cdisc.org/pages/viewpage.action?pageId=157583203

cdisc.

External Communications

- COSA Spotlight Webinars
- Interchange presentations
- Social media posts
- Projects included in Dataset-JSON presentations
- Tools available for use in future pilots

cdisc.

Thank You!

Questions?

Sam Hume
shume@cdisc.org
https://www.linkedin.com/in/sam-hume-dsc

Charles Shadle
cshadle@cdisc.org