



MLQP Card Generator

Guided Qualification Cards for ML Dataset and Model TEVV

ARiA

Haley Green
haley.green@ariacoustics.com



AI4SE & SE4AI Workshop 2026

Introduction - The ML V&V Challenge

AI/ML is increasingly augmenting decision-making and reducing workload as a core component for safety-critical systems.

Challenges:

1. The safety, suitability, and supportability of ML systems rely on domain-matched training data over traditional software V&V testing.
2. Rapid ML retraining and qualification cycles render traditional software assurance processes untenable.

Work presented is in support of STTR-A22B-T002, Verification, Validation, Assurance and Trust of AI/ML Models and Data for Safety-Critical Systems, with DEVCOM Armaments Center

Introduction - MLSSEG Foundation

MLSSEG (supporting MIL-STD-882E) establishes essential safety engineering practices for ML.

Provides comprehensive traceability and standardization across the ML lifecycle.

Challenges:

1. Practitioners must independently interpret these requirements, which **introduces variability in how they are applied.**
2. Inconsistent interpretations limit cross-discipline collaborations, **slowing a workflow that demands efficiency.**

**Machine Learning System Safety Engineering Guide
with Implementation Process and Tasks
Supporting MIL-STD-882E w/CHANGE 1**



January 2026

Office of Systems Engineering and Architecture

Office of the Under Secretary of War
for Research and Engineering

Washington, D.C.

DISTRIBUTION STATEMENT A. Approved for public release. Distribution is unlimited.

Proposed Solution - The MLQP Card Generator



Translates complex Level of Rigor (LOR) tasks into intuitive, practitioner-facing questions.



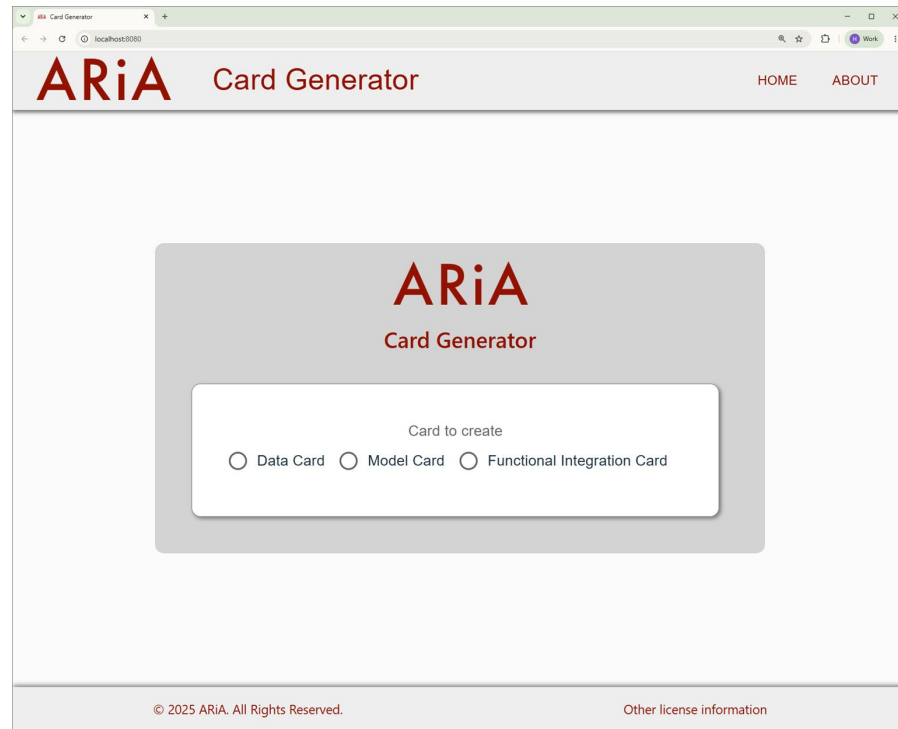
Guides users through Data, Model and Functional Integration Card completion via an interactive UI.



Automatically generates human-readable reports mapped directly to relevant LORs.

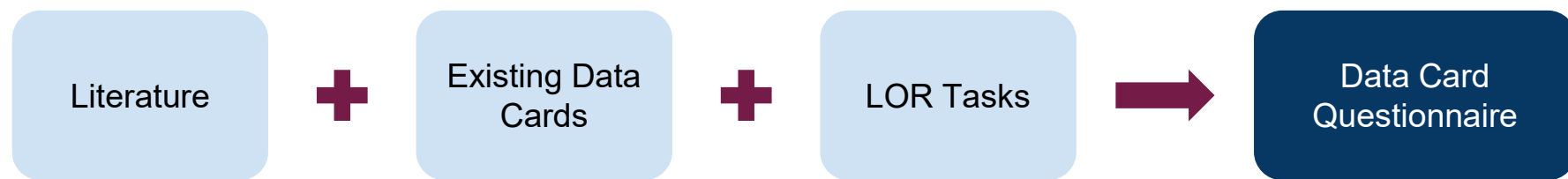


Establishes consistent documentation practices that reduce interpretation time.



Methodology - Card Development

Initial Approach



- Unrealistic → too many questions
- Asked users to self-assess (e.g., “Is there leakage?”)

Key Realization: Fully completed simpler card > partially completed thorough card.

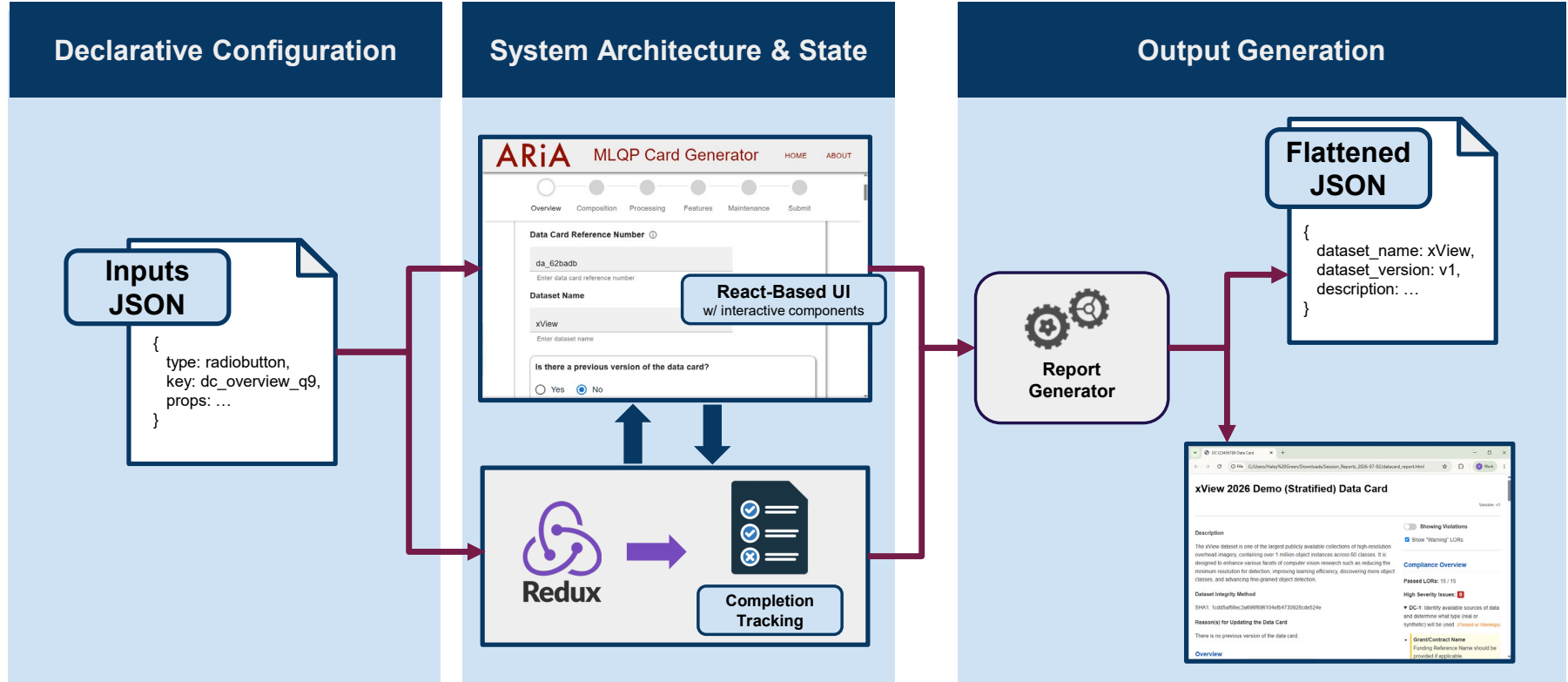
Methodology - Card Development

Changes Made

- Narrowed to image classification and object detection
- Reduced open-ended questions (by leveraging checkboxes, dropdowns, radio buttons)
- Nested questions
- Mapped all questions directly to LORs

Section	Question #	Question Text	Follow-Up?	Type	LOR
Selection	50	Which best describes the algorithm used?	51, 52	Radio	RA-2, AL-2
Selection	51	Is this algorithm widely used in practice or in literature for this type of task?		Yes/No	RA-2, AL-1
Selection	52	Why was this algorithm selected instead of more common alternatives?		Text	RA-2, AL-1
Selection	53	Which backbone does your model use?		Text	DA-7, AL-3
Selection	54	Which neck/feature aggregation network does your model use?		Text	DA-7, AL-3
Selection	55	Which head/task layer does your model use?		Text	DA-7, AL-3
Selection	56	Was an autoencoder used in the development of the model?	57	Yes/No	AL-3

Methodology - System Implementation



Achieved Outcomes - The Cards



Data Card

Handles dataset lineage, collection methodologies, and feature identification.



Model Card

Documents model architecture, training hyperparameters, and anticipated performance bounds.



Functional Integration Card

Functional integration (FI) into system practice and preventative failure scenario documentation.

Data Card

Section	Description
Overview	Basic metadata (collection location, data type)
Composition	Makeup (how data was generated/V&V)
Processing	Loading, extraction, labeling information
Features	Key attributes in the dataset for classification/detection
Maintenance	Maintenance, use, and distribution documentation
Submit	Summary of the data card answers

The screenshot shows a web browser window with the title "ARiA Card Generator". The browser's address bar shows "localhost:3030". The page has a header with the "ARiA Card Generator" logo and navigation links for "HOME" and "ABOUT". The main content area features a large gray box with the "ARiA Card Generator" logo and a white box titled "Card to create". This box contains three radio button options: "Data Card", "Model Card", and "Functional Integration Card". The "Data Card" option is selected. At the bottom of the page, there is a footer with the text "© 2025 ARiA. All Rights Reserved." and a link for "Other license information".

Model Card

Section	Description
Overview	Basic metadata (storage location, operational task)
Selection	model selection process and anticipated performance
Maintenance	Maintenance, use, and distribution documentation
Submit	Summary of the model card answers

The screenshot shows a web browser window with the title 'ARiA Card Generator'. The URL bar shows 'localhost:3030'. The page has a header with the 'ARiA Card Generator' logo and navigation links for 'HOME' and 'ABOUT'. The main content area features a large gray box with the 'ARiA Card Generator' logo and a white box containing the text 'Card to create' and three radio button options: 'Data Card', 'Model Card', and 'Functional Integration Card'. The footer contains the copyright notice '© 2025 ARiA. All Rights Reserved.' and a link for 'Other license information'.

Functional Integration (FI) Card

Section	Description
Overview	Basic overview of the FI (SFCI, backup safeguards)
Assessment	Potential failure scenarios and planned response
HMT	Document how humans will interact with the model
Mitigation	Countermeasures for intervention and primary contacts
Submit	Summary of the FI card answers

The screenshot shows a web browser window displaying the ARiA Card Generator application. The browser's address bar shows the URL as localhost:3030. The application has a header with the ARiA logo and the text 'Card Generator', along with navigation links for 'HOME' and 'ABOUT'. The main content area features a large gray box with the ARiA logo and 'Card Generator' text. Below this, a white box contains the heading 'Card to create' and three radio button options: 'Data Card', 'Model Card', and 'Functional Integration Card'. The 'Functional Integration Card' option is currently selected. At the bottom of the page, there is a footer with the text '© 2025 ARiA. All Rights Reserved.' and a link for 'Other license information'.

Achieved Outcomes - Card Report

Compliance Tracking

Tracks unsatisfied requirements, with a checkbox to filter out warnings.

Interactive Traceability

Hovering over form fields displays the LORs tied to that specific response.

Visual Mapping

Color-codes fields by LOR category to visually map questions to requirements.

Distribution Friendly

Interactive HTML, JSON, and PDF versions to support distribution.

xView 2026 Demo (Stratified) Data Card

Version: v1

Description

The xView dataset is one of the largest publicly available collections of high-resolution overhead imagery, containing over 1 million object instances across 60 classes. It is designed to enhance various facets of computer vision research such as reducing the minimum resolution for detection, improving learning efficiency, discovering more object classes, and advancing fine-grained object detection.

Dataset Integrity Method

SHA256: c46e97b2a6506f9be596ae0991c7058b0c0e70a73a7b0d4688b25829b0f2d09f

Reason(s) for Updating the Data Card

There is no previous version of the data card.

Compliance Overview

Showing Violations: ☒ Show "Warning" LORs

Passed LORs: 12 / 15

High Severity Issues: 1

- DC-1. Identify available sources of data and determine what type (real or synthetic) will be used
- Does your dataset include data collected from the actual operational environment? Dataset should include operational data for compliance. *Provided: No*
- Grant/Contract Name Funding Reference Name should be provided if applicable. *Provided: N/A*
- Grant/Contract Number Funding Reference Number should be provided if applicable. *Provided: N/A*
- Was any of the data synthetically generated? Must indicate if synthetic data was used. *Provided: None*
- DC-3. Document the type of data and sources/conditions/metadata associated with it
- Does your dataset include data collected from the actual operational environment? Dataset should include operational data for compliance. *Provided: No*

Overview

Data Storage

Stored Locally in /home/muser/xview_demo/xview_split_stratified

Data Format

- Images/video
- JSON

Dataset Funding

Grant: *N/A*

Ref No: *N/A*

Data Collection

Start Date: Wed, 13 Aug 2014 04:00:00 GMT

End Date: Mon, 01 Jan 2018 05:00:00 GMT

Collection Regions

Dataset Author

Dariusz Lam, Defense Innovation Unit Experimental (DIUx)

Phone: *N/A*

Email: *N/A*

Dataset Sponsor

DIUx and the National Geospatial-Intelligence Agency (NGA)

Intended User

Research and Academic Community

**This dataset is available for free public use*

Ongoing Work

Refining the Cards

- Adding tooltips
- Autofill ingested data
- Test extensively with operational datasets

The MLQP Dashboard

Tool-based metrics (DataEval, ModelEval, NRTK) used alongside user-defined card entries for TEVV.

