

Here is the checklist for the author developing the Artificial Intelligence - Machine Learning software and the reviewer reviewing it:

DQRs (Data Quality Requirements) Checklist

Training Data Set

- Is the training data set relevant to support the intended use of the AI/ML software? (yes/no)
- Can the origin of the training data be determined? (yes/no)
- Are there requirements related to the annotation process for the training data? (yes/no)
- What is the format, accuracy, and resolution of the training data? (specify)
- Is the training data set complete and representative? (yes/no)
- Are the training, validation, and test data sets independent? (yes/no)

Data Accuracy

- What is the accuracy of the training data? (specify)
- What is the resolution of the training data? (specify)

Annotated Data Quality

- What is the quality of the annotated training data? (good, fair, poor)
- Are there any necessary manipulations of the training data (e.g. anonymization)? (yes/no)

Integrity and Corruption Prevention

- Has the training data been corrupted while stored, processed, or transmitted over a communication network? (yes/no)
- Are there mechanisms in place to prevent corruption of the training data? (yes/no)

Data Representativeness

- Is the training data representative of the intended use case? (yes/no)
- Are there any biases or anomalies present in the training data? (yes/no)

Reviewer's Checklist

- Have all DQRs been met for the training, validation, and test data sets? (yes/no)
- Is the data quality sufficient to support the intended use of the AI/ML software? (yes/no)
- Are there any gaps or inconsistencies in the data that could affect the performance of the AI/ML software? (yes/no)

Note: This checklist is a deep dive into the requirements for Data Quality Requirements (DQRs) and is intended to be used by both the author developing the AI/ML software and the

reviewer reviewing it.