

Here is the checklist in markdown format:

Data Management Checklist

Data Collection and Preparation

Question	Yes/No
Has data been collected for training, testing, and verification?	?
Are all necessary data sets captured (e.g. nominal data, edge cases, corner cases)?	?
Have infeasible corner cases and novelties been detected and removed?	?
Are outliers defined and managed?	?

Data Labelling (Supervised Learning)

Question	Yes/No
Has data been labelled correctly for supervised learning?	?
Is the annotation process well-defined and traceable?	?

Data Set Verification

Question	Yes/ No
Are the training, validation, and test data sets independent of each other?	?
Have the completeness, representativeness, and accuracy of the data sets been verified?	?
Is the integrity of the data ensured (i.e. no corruption during storage or transmission)?	?

Data Management Requirements

Question	Yes/ No
Are the necessary manipulations of the data (e.g. anonymization) defined and performed?	?
Has the origin of the data been determined and documented?	?

Derived Requirements

Question	Yes/No
Have derived requirements from a safety perspective been reviewed and validated?	?

I hope this helps! Let me know if you have any further requests. 😊