

## **Checklist for Developing AI/ML Software**

### **Identity and Purpose**

- Has the definition of AI/ML constituent ODD parameters been established based on relevant industry standards? (Yes/No)
- Have nominal data definitions been provided? (Yes/No)
- Have edge cases, corner cases, and infeasible corner cases been identified for stability of the model? (Yes/No)
- Have outliers been defined for detection and management? (Yes/No)
- Are there mechanisms to detect and remove inliers? (Yes/No)
- Are there mechanisms to detect and manage novelties? (Yes/No)

### **Data Quality Requirements**

- Has the relevance of data been established to support intended use? (Yes/No)
- Can the origin of the data be determined? (Yes/No)
- Have requirements for annotation process been defined? (Yes/No)
- Are the format, accuracy, and resolution of the data specified? (Yes/No)
- Is there a mechanism to ensure traceability of data from origin to final operation? (Yes/No)
- Is there a mechanism to prevent corruption during storage, processing, or transmission? (Yes/No)
- Is the completeness and representativeness of data sets ensured? (Yes/No)
- Are the training, validation, and test data sets independent? (Yes/No)

### **Data Management**

- Have data management requirements been captured? (Yes/No)
- Has data collection been performed? (Yes/No)
- Is data labelling (supervised learning) done? (Yes/No)
- Has data preparation (pre-processing, transformation, feature engineering) been conducted? (Yes/No)
- Are the various data sets used in the learning phase identified? (Yes/No)
- Have data set verifications been performed (accuracy, completeness, representativeness)? (Yes/No)
- Are there independence requirements between data sets? (Yes/No)
- Has unwanted bias inherent to the data sets been identified and eliminated? (Yes/No)

Note: This checklist is designed to be low-level and deep down in the requirement, focusing on specific aspects of AI/ML software development.