

Here is the generated checklist for the author developing the artificial intelligent-machine learning software and for the reviewer reviewing it:

## **Data Management**

### **Data Management Requirements Capture**

<b>Requirement</b>	<b>Yes/No</b>
Is the data management process clearly defined?	
Are the objectives to be satisfied at each individual data point identified?	
Are the data sets used in the project development lifecycle identified?	
Are the various data set types (e.g. training, validation, test) distinguished?	

### **Data Collection**

<b>Requirement</b>	<b>Yes/No</b>
Is the source of each data set clearly identified?	
Are data quality issues addressed?	
Are data inconsistencies or errors handled?	
Are data sets verified for accuracy and completeness?	

### **Data Labelling (only in supervised learning)**

<b>Requirement</b>	<b>Yes/No</b>
Is the labelling process consistent across all data sets?	
Are labels accurate and reliable?	
Are data sets labelled according to the ML requirements?	

### **Data Preparation (pre-processing, data transformation, feature engineering)**

<b>Requirement</b>	<b>Yes/No</b>
Are data sets pre-processed for quality and consistency?	
Are features engineered or transformed as needed?	
Is data prepared for use in the ML model?	

### **Data Set Verification**

<b>Requirement</b>	<b>Yes/No</b>
Is each data set verified for accuracy, completeness, and representativeness?	
Are data sets checked against the ML requirements?	
Are data sets verified with respect to the AI/ML constituent ODD?	

### **Independence Requirements between Data Sets**