Provenance: A Machine Learning Non-Functional Requirement

Full Requirement Specifications List

Lynn Vonderhaar, Timothy Elvira, Omar Ochoa

*[Req. 1] The model’s pipeline shall incorporate prediction explanations for the user.*

*[Req. 2] The model’s behavior development shall be transparent to the user.*

*[Req. 3] The model’s development artifacts shall be traceable throughout its lifecycle to ensure transparency.*

*[Req. 4] The model’s development pipeline shall save data provenance to ensure development artifact traceability to model requirement specifications.*

[Req. 5] The model’s development pipeline shall save the method of data collection to the provenance documentation.

[Req. 6] The model’s development pipeline shall save the date of data collection to the provenance documentation.

[Req. 7] The model’s development pipeline shall save the name of the person collecting the data to the provenance documentation.

[Req. 8] The model’s development pipeline shall save the data transformation steps to the provenance documentation.

[Req. 9] The model’s development pipeline shall save the date of the data transformation steps to the provenance documentation.

[Req. 10] The model’s development pipeline shall save the name of the developer performing the data transformation steps to the provenance documentation.

*[Req. 11] The model’s development pipeline shall save model provenance to ensure development artifact traceability to model requirement specifications.*

[Req. 12] The model’s development pipeline shall save model versions to the provenance documentation.

*[Req. 13] The model’s development pipeline shall save maintenance and revision provenance to ensure development artifact traceability to model requirement specifications.*

[Req. 14] The model’s development pipeline shall save requirement changes to the provenance documentation.

*[Req. x] The model’s behavior shall be ethically sound.*

*[Req. x]The model’s predictions shall be justifiably correct.*

*[Req. x] The model’s behavior shall not show signs of bias.*

*[Req. x] The data provenance shall not show signs of bias.*

[Req. x] The model’s development pipeline shall save development data design decisions made to protect against bias to the provenance documentation.

[Req. x] The model’s development pipeline shall save the model’s training data to the provenance documentation.

[Req. x] The model’s development pipeline shall save the training data’s class proportions to the provenance documentation.

[Req. x] The model’s development pipeline shall save the validation data’s class proportions to the provenance documentation.

[Req. x] The model’s development pipeline shall save the testing data’s class proportions to the provenance documentation.

[Req. x] The model’s development pipeline shall save the training data’s annotations to the provenance documentation.

[Req. x] The model’s development pipeline shall save the validation data’s annotations to the provenance documentation.

[Req. x] The model’s development pipeline shall save the testing data’s annotations to the provenance documentation.

[Req. x] The model’s development pipeline shall save the date of the training data’s annotations to the provenance documentation.

[Req. x] The model’s development pipeline shall save the date of the validation data’s annotations to the provenance documentation.

[Req. x] The model’s development pipeline shall save the date of the testing data’s annotations to the provenance documentation.

[Req. x] The model’s development pipeline shall save the name of the training data’s annotator to the provenance documentation.

[Req. x] The model’s development pipeline shall save the name of the validation data’s annotator to the provenance documentation.

[Req. x] The model’s development pipeline shall save the name of the testing data’s annotator to the provenance documentation.

*[Req. x] The verification and behavior provenance shall not show signs of bias.*

[Req. x] The model’s development pipeline shall save the model’s achieved performance metrics to the provenance documentation.

[Req. x] The model’s development pipeline shall save the model’s synthetic dataset verification results to the provenance documentation.

[Req. x] The model’s development pipeline shall save the model’s testing history to the provenance documentation.

*[Req. x] The deployment provenance shall not show signs of bias.*

[Req. x] The model’s development pipeline shall save the model’s prediction logs to the provenance documentation.

*[Req. x] The model’s behavior shall be maintainable.*

*[Req. x] The model’s pipeline shall save model provenance to help trace behavior.*

[Req. x] The model’s development pipeline shall save the model version to the provenance documentation.

[Req. x] The model’s development pipeline shall save the model’s training data to the provenance documentation.

[Req. x] The model’s development pipeline shall save the training data’s class proportions to the provenance documentation.

[Req. x] The model’s development pipeline shall save the validation data’s class proportions to the provenance documentation.

[Req. x] The model’s development pipeline shall save the testing data’s class proportions to the provenance documentation.

[Req. x] The model’s development pipeline shall save the training data’s labels to the provenance documentation.

[Req. x] The model’s development pipeline shall save the pretrained model’s provenance to the provenance documentation.

[Req. x] The model’s development pipeline shall save the requirement prioritization to the provenance documentation.

[Req. x] The model’s development pipeline shall save the data inclusion and exclusion criteria to the provenance documentation.

*[Req. x] The model’s behavior shall be adaptable to a new domain.*

*[Req. x] The model’s development pipeline shall save model provenance to trace behavior.*

[Req. x] The model’s development pipeline shall save the model version to the provenance documentation.

[Req. x] The model’s development pipeline shall save the model’s training data to the provenance documentation.

[Req. x] The model’s development pipeline shall save the training data’s class proportions to the provenance documentation.

[Req. x] The model’s development pipeline shall save the validation data’s class proportions to the provenance documentation.

[Req. x] The model’s development pipeline shall save the testing data’s class proportions to the provenance documentation.

[Req. x] The model’s development pipeline shall save the training data’s labels to the provenance documentation.

[Req. x] The model’s development pipeline shall save the pretrained model’s provenance to the provenance documentation.

[Req. x] The model’s development pipeline shall save the requirement prioritization to the provenance documentation.

[Req. x] The model’s development pipeline shall save the data inclusion and exclusion criteria to the provenance documentation.

*[Req. x] The model’s behavior shall be interpretable based on the training data.*

*[Req. x] The model’s development pipeline shall save data provenance to trace behavior.*

[Req. x] The model’s development pipeline shall save the training dataset to the provenance documentation.

[Req. x] The model’s development pipeline shall save the data source to the provenance documentation.

[Req. x] The model’s development pipeline shall save the data fidelity to the provenance documentation.

[Req. x] The model’s development pipeline shall save the model’s parameters to the provenance documentation.

[Req. x] The model’s development pipeline shall save the model’s hyperparameters to the provenance documentation.

[Req. x] The model’s development pipeline shall save the dataset versions to the provenance documentation.

[Req. x] The model’s development pipeline shall save the model versions to the provenance documentation.

*[Req. x] The model’s behavior shall be testable.*

*[Req. x] The model’s behavior shall be tested using provenance from the requirements phase.*

[Req. x] The model’s development pipeline shall save the customer requirement artifacts to the provenance documentation.

[Req. x] The model’s development pipeline shall save the model quality attributes to the provenance documentation.

[Req. x] The model’s development pipeline shall save the requirement prioritization to the provenance documentation.

[Req. x] The model’s development pipeline shall save the requirement specifications to the provenance documentation.

[Req. x] The model’s development pipeline shall save the constraints to the provenance documentation.

[Req. x] The model’s development pipeline shall save the customer requirement artifacts to the provenance documentation.

*[Req. x] The model’s behavior shall be tested using provenance from the model phase.*

[Req. x] The model’s development pipeline shall save the model speed to the provenance documentation.

[Req. x] The model’s development pipeline shall save the model size to the provenance documentation.

*[Req. x] The model’s behavior shall be tested using provenance from the verification and behavior phase.*

[Req. x] The model’s development pipeline shall save the performance metrics to the provenance documentation.

*[Req. x] The model’s behavior shall be tested using provenance from the deployment phase.*

[Req. x] The model’s development pipeline shall save the prediction logs to the provenance documentation.

*[Req. x] The model’s behavior shall be reproducible.*

*[Req. x] The model’s development pipeline shall save requirement provenance for model behavior reproduction.*

[Req. x] The model’s development pipeline shall save the requirement specifications to the provenance documentation.

[Req. x] The model’s development pipeline shall save the constraints to the provenance documentation.

*[Req. x] The model’s development pipeline shall save data provenance for model behavior reproduction.*

[Req. x] The model’s development pipeline shall save the datasets to the provenance documentation.

[Req. x] The model’s development pipeline shall save the data type to the provenance documentation.

[Req. x] The model’s development pipeline shall save the data source to the provenance documentation.

[Req. x] The model’s development pipeline shall save the data fidelity to the provenance documentation.

[Req. x] The model’s development pipeline shall save the data metadata to the provenance documentation.

[Req. x] The model’s development pipeline shall save the synthetic data metadata to the provenance documentation.

[Req. x] The model’s development pipeline shall save the method of collection to the provenance documentation.

[Req. x] The model’s development pipeline shall save the inclusion and exclusion criteria to the provenance documentation.

[Req. x] The model’s development pipeline shall save the preprocessing steps to the provenance documentation.

[Req. x] The model’s development pipeline shall save the labels to the provenance documentation.

[Req. x] The model’s development pipeline shall save the feature selection to the provenance documentation.

[Req. x] The model’s development pipeline shall save the training, validation, and testing splits to the provenance documentation.

[Req. x] The model’s development pipeline shall save the training, validation, and testing class proportions to the provenance documentation.

*[Req. x] The model’s development pipeline shall save model provenance for model behavior reproduction.*

[Req. x] The model’s development pipeline shall save the pretraining provenance to the provenance documentation.

[Req. x] The model’s development pipeline shall save the model architecture to the provenance documentation.

[Req. x] The model’s development pipeline shall save the parameters to the provenance documentation.

[Req. x] The model’s development pipeline shall save the hyperparameters to the provenance documentation.

[Req. x] The model’s development pipeline shall save the optimization algorithms to the provenance documentation.

[Req. x] The model’s development pipeline shall save the environment configurations to the provenance documentation.

[Req. x] The model’s development pipeline shall save the source code to the provenance documentation.

*[Req. x] The model’s development pipeline shall save verification and behavior provenance for model behavior reproduction.*

[Req. x] The model’s development pipeline shall save the performance metrics to the provenance documentation.

[Req. x] The model’s development pipeline shall save the testing history to the provenance documentation.

[Req. x] The model’s development pipeline shall save the model versions to the provenance documentation.

*[Req. x] The model’s development pipeline shall save deployment provenance for model behavior reproduction.*

[Req. x] The model’s development pipeline shall save the operation location to the provenance documentation.

[Req. x] The model’s development pipeline shall save the licenses to the provenance documentation.

[Req. x] The model’s development pipeline shall save the containers to the provenance documentation.