Trustworthy AI Procurement Card™

Trustworthy AI Procurement Card™ is a non-exhaustive list of information which can accompany acquisition decisions.[1] The Card is similar to the DataSheets or Model Cards, in the sense that the objective is to promote transparency and better due diligence during AI procurement process. Buyers could require the vendors to provide this information during the process. The Card can serve as a quick reference for decision-makers, enabling easy comparison between vendor models, as well as comparison between updates to the same product.

The Card can be used as a self-conformity assessment tool first, where the vendor provides the details of the product to the buyer, and documents its decisions and practices.

In cases where the procurement action is to decide on the development of a product, the Card can be used to establish the statement of work.

The buyer can use the Card as a baseline for its internal Test, Evaluation, Verification and Validation (TEVV) process.

And finally, the Card can contribute to the development and maintenance of AI use case registries where information critical for public awareness and accountability is provided.

Category	Documentation
Intended purpose	Quantitative and qualitative analysis to demonstrate the expected benefit
Objective function	o Explanation of the process to determine objective function o Assumptions included in the objective function
Datasets	o Process for feature/variable selection o Information on data provenance, and limitations of data, and where applicable labeling procedures o Main data processing decisions o Demonstration of dataset quality
Models	o Justification for selected approach o General logic of the system o Processes for model training, development, fine-tuning, and validation
Environment	Hardware requirements for development, hosting, and operational use
Solution performance	o Justification for selected performance metrics and thresholds o Embedded trade-offs between performance metrics o Outlier evaluation process o Comparison to baseline existing method
Vendor performance	Diversity of teams' experience, background, and domain expertise
Dependencies	Upstream and downstream connections to other data flows and AI systems
Validity	Documentation of external and internal validity, and scientific evidence of construct validity
Risk assessments	o Possible risks and harms identified by the vendor o Relevant mitigation strategies o Methods for continuous monitoring (to safeguard against performance changes and drift)
Safety and security	o Processes to ensure robustness and resiliency o Controls against adversarial attacks o Embedded privacy design decisions o Quality assurance for third-party code libraries and APIs
Human oversight	o Necessary user training and measures for meaningful oversight o Fallback mechanisms/Plans
Technical elements of Trustworthy AI	Provide metrics or explanation of the elements the solution incorporates or facilitates: Accuracy, Explainability, Interpretability, Traceability, Reliability, Reproducibility, Contestability

[1] Trustworthy AI Procurement Card™ is developed by Merve Hickok.