ERA SCIENCES
Evolved DI Solutions

Model Details

Overview
• Description of the model, including inputs and outputs
• Type of learning approach (supervised, unsupervised, or reinforcement).
• Whether the model employs static (fixed after training) or dynamic (continuously learning)
• Model architecture and structure, including frameworks/libraries used and hardware/software dependencies

GxP Justification
• Justification for GxP applicability based on system use (e.g., clinical trial analysis, product quality monitoring).

Version
• Model versioning information, including major and minor updates, date of release and summary of changes between versions.

Owners
• Business owner responsible for decision-making and compliance.
• System owner responsible for maintenance and updates.

References
• Any further internal or external references

Considerations

Intended use
• Specific business application (e.g., visual inspection, quality control, patient outcome prediction).
• Operational scope and limitations.

Caveats & limitations
• Model constraints (e.g., performance thresholds, edge cases).
• Implications of false positives and false negatives in a regulated setting.
• Limitations in knowledge representation ("the model doesn't know what it doesn't know").

Warnings & residual risks
• Risks associated with incorrect predictions.
• Summary of residual risks identified in risk assessment.
• Known failure cases and strategies for mitigation.
•Mechanisms for handling performance deviations and retraining.

Ethical considerations
• Use of sensitive or personally identifiable data.
• Subgroup fairness and necessary rebalancing.
• Explainability of model decisions for regulatory transparency.

Further recommendations
• Recommendations for model monitoring and governance.
• Periodic review schedule and performance revalidation.
• Change management process for updates and modifications.

Quantitative analysis

Evaluation metrics
• Statistical performance benchmarked against traditional (non-ML) methods.
• Statistical error rates (confidence intervals, p-values).
• Report confusion matrix (false positives, false negatives).
• Explain trade-offs between precision and recall.
• Robustness and generalizability testing across different datasets.
• Stability testing to assess drift over time.

Data detail
• Data sources and origin, including regulatory considerations (e.g., PHI, PII, proprietary clinical data).
• Data pre-processing steps (normalization, tokenization, de-identification).
• Dataset composition and split (training, validation, test).
• Management of outliers and anomalous data.
• Data labeling: manual, automated, or hybrid labeling.
• Data labeling: labeling authority (subject matter experts, crowdsourcing).
• Bias mitigation: Representation fairness (e.g., demographic balance).
• Bias mitigation: Clinical trial diversity considerations.