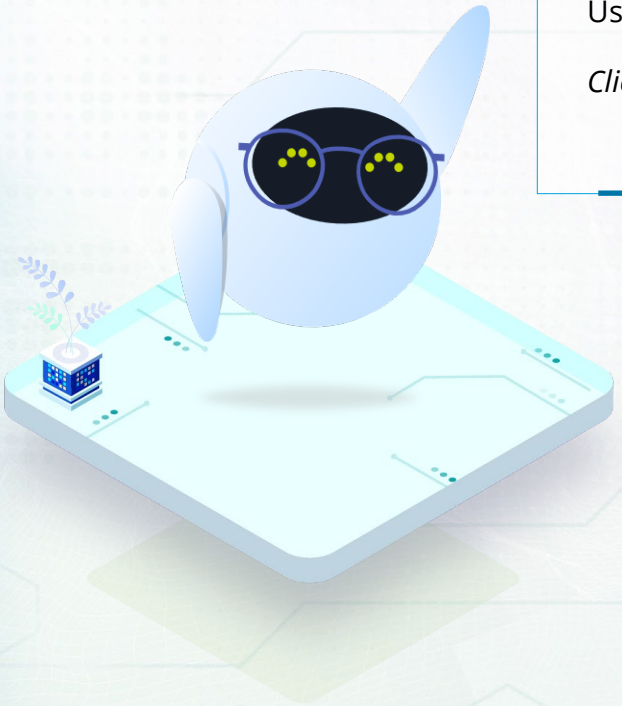
Deloitte.

Trustworthy AI™

Module 2: Tools Ecosystem



Use this document to refer to the Trustworthy AI™ tools ecosystem.

Click each tab to view the list of AI tools.

Deloitte's Proven AI Tools

Deloitte Access to Hyperscaler Tools

Deloitte Access to MLOps Tools

Deloitte's Proven AI Tools

Deloitte's AI accelerators allow access to a number of toolkits and frameworks that enable the establishment of effective control structures, AI/ML infrastructure, and algorithms that enable reliable machine learning (ML) models.

Example Tools	Description
Control Structure Framework	The Controls Structure Framework is an AI controls inventory and AI technique library that can be used to establish an effective AI control structure, while working collaboratively with client data science and risk management teams to help mitigate risk.
D-Bias	Deloitte's in-house toolkit, D-Bias, was developed to conduct automated code checks, and detect & prescriptively correct bias errors and their underlying sources. It helps identify general bias in data and discover potential biases existing in the functional form of the algorithm.
GlassBox	GlassBox is designed to monitor, interpret and scrutinize AI algorithms. The solution offers an end-to-end platform which can be embedded within the organization. Furthermore, the solution provides embedded definitions of fairness and bias detection, assessed with statistical thresholds. The approach can be used in, among others, continuous monitoring process, or provide explanation of AI decisions on demand (for example, to address GDPR requirements).
Fusion Architecture for Learning & Cognition ("FALCON")	FALCON is a three-field, two-layer neural network model based on adaptive resonance theory (ART) for driving explainability in AI models. It learns incrementally and in real-time, using competitive and memory-based learning paradigms to facilitate algorithm enhancement and explainability by identifying key predictors and redundant data.

Deloitte's Proven AI Tools

Deloitte's AI accelerators allow access to a number of toolkits and frameworks that enable the establishment of effective control structures, AI/ML infrastructure, and algorithms that enable reliable machine learning (ML) models.

Example Tools	Description
Model Guardian	Guardian is an end-to-end, customizable algorithmic fairness and bias testing tool to identify, investigate, and track biases in credit risk models.
Lucid	Lucid is a business intelligence platform which provides insights and applies causal reasoning to complex business problems.
IBM AI Fairness 360	IBM's AI Fairness 360 (AIF360) is a toolkit of nine algorithms focused on examining, reporting, and mitigating discrimination and bias in ML models. The open-source toolkit is available as a Python API, a Python package, and an R package.
Fairlearn	Microsoft's Fairlearn python library incorporates interactive visualization and bias assessment, then applies algorithms to mitigate bias that is uncovered. The tool focuses on quality of service and allocation metrics across groups (e.g. race, gender, age, disability status).
Manifold	Manifold by Uber illustrates predictor importance for black box algorithms, over the entire population as well as across segments identified through clustering.
What-If Tool	Google's What-If Tool focuses on AI fairness and transparency, provides visualization of model performance, supports what-if analyses, and assesses algorithmic fairness across five measures.

Deloitte Access to Hyperscaler Tools

With access to over a dozen hyperscaler tools, their usage enables improved industry practices and delivering value such as: reducing bias, increasing explainability, and improving consumer and privacy data protection.

Example Tools	Description
Chatterbox (AIMI)	Chatterbox Labs AI Model Insights (AIMI) is a technology solution to explain, trace, action, score bias, test, detect weaknesses, imitate and identify privacy risks in enterprise AI models.
DataRobot	DataRobot is an Augmented Intelligence platform that aims to enable AI-driven enterprises by deploying end-to-end, automated best practices in the form of built-in guardrails, model documentation, prediction warnings, and humility rules.
Darwin AI GENSYTH	GENSYNTH by DarwinAI allows data and advanced factory automation to be leveraged to assist in cases such as quality inspection. It can be utilized for defect detection to improve quality control, and it enables an increase in output and efficiency via 'adaptive factory automation.'
Google Deepmind	Google Deepmind offers general-purpose learning algorithms that can be utilized for healthcare applications (i.e., disease diagnosis, medical scan analysis).
Google Explainable AI	Googles Explainable AI increases the explainability of AI by providing reasoning surrounding predictions, which enables bias to be reduced and an overall increase to human 'readability.'
Google Differential Privacy	Google's Differential Privacy tool is utilized for privacy and consumer data protection. The tool allows to derive insights from data while simultaneously ensuring the privacy of individuals in the dataset.
Element AI Knowledge Scout	Element AI Knowledge Scout enables information gathering workflows by integrating knowledge graphs with question-answering, and semantic search capabilities into a unified enterprise software experience to help users navigate data.
Element AI XAI	XAI develops models that are easier to understand, namely, model explainability. Alternatively, XAI extracts explanations from complex pre-developed models that are otherwise difficult to understand for their users, namely, post-hoc explainability.

Deloitte Access to Hyperscaler Tools

With access to over a dozen hyperscaler tools, their usage enables improved industry practices and delivering value such as: reducing bias, increasing explainability, and improving consumer and privacy data protection.

Example Tools	Description
MIT CSAIL	MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL) conducts research in areas related to computer science and AI, such as robotics, systems, theory, biology, machine learning, speech recognition, vision and graphics.
IBM Watson OpenScale	IBM Watson OpenScale helps organizations detect and correct AI model bias and drift, explain AI outcomes, monitor model accuracy, and analyze payloads,
Microsoft MS SEAL	Microsoft SEAL is a homomorphic encryption library that allows additions and multiplications to be performed on encrypted integers or real numbers.
Microsoft Interpret ML	Interpret ML is an open-source package that incorporates machine learning interpretability techniques to train interpretable Glassbox models and explain blackbox systems.
IBM Adversarial Robustness Toolbox	Adversarial Robustness Toolbox (ART) is a Python library for Machine Learning Security. ART provides tools that enable developers and researchers to defend and evaluate Machine Learning models and applications against the adversarial threats of Evasion, Poisoning, Extraction, and Inference.
IBM HElib	HElib is an open-source software library that implements homomorphic encryption (HE). The library implements the Brakerski-Gentry-Vaikuntanathan (BGV) homomorphic encryption scheme, as well as optimizations such as Smart-Vercauteren ciphertext packing techniques.

Deloitte Access to Machine Learning Operations Tools

With over ten machine learning operations tools, Deloitte has access to platforms and enterprise scales that ease the creation and development of machine learning models.

Example Tools	Description
Algorithmia	Algorithmia automates and accelerates delivery of ML models into production, driving a process of continuous optimization across all stages of the ML lifecyle within existing operational processes.
Amazon SageMaker	Amazon SageMaker enables developers and data scientists to build, train, and deploy machine learning (ML) models at scale. As an ML service, it provides tools for each step of the ML development, including labeling, data preparation, feature engineering, statistical bias detection, auto-ML, training, tuning, hosting, explainability, monitoring, and workflows.
Databricks MLFlow	Databricks MLFlow is an open-source platform for managing the end-to-end machine learning lifecycle. It enables tracking, deployments and management of models, packaging of ML code, and a centralized model store for managing models' full lifecycle stage transitions: from staging to production, with capabilities for versioning and annotating.
Datatron	Datatron provides an enterprise scale automated model monitoring and model governance platform for ML, AI, and Data Science models in production.
Domino	Domino's enterprise MLOps platform accelerates research, speeds model deployment, and increases collaboration for code-first data science teams at scale.
Google Kubeflow	Google's Kubeflow is an open-source machine learning platform that deploys and manages end-to-end ML workflows. There are additional tools to support complex workflows running on Kubernetes.

Deloitte Access to Machine Learning Operations Tools

With over ten machine learning operations tools, Deloitte has access to platforms and enterprise scales that ease the creation and development of machine learning models.

Example Tools	Description
H20	H2O MLOps is a system that enables the deployment, management, and governance of models in production with seamless integration to H2O Driverless AI and H2O open source for model experimentation and training. It supports model management by providing tools such as A/B testing and automatic retraining.
Microsoft Azure	Microsoft Azure MLOps tools allow for training reproducibility with advanced tracking of datasets and code in addition to autoscaling for model training and deployment. It also supports efficient workflows with scheduling and management capabilities to build and deploy with continuous integration/continuous deployment (CI/CD).
SAS Model Manager	SAS Model Manager allows models to be stored within folders or projects, candidate models to be developed and validated, and enables the assessment of candidate models for champion model selection. In addition, the analytic lifecycle can be automated via custom workflows.
Weights & Biases	Weights & Biases offers several visualization tools for machine learning models to allow model performance visualizations and provide automation of AI model training support.

Deloitte.