

# Comprehensive Healthcare AI Glossary

---

**Accuracy:** The degree to which the output of an AI model correctly reflects the true values—critical in clinical predictions and diagnostics.

**Algorithm:** A set of instructions that guides how a machine learns or makes decisions from data. In healthcare, algorithms are the backbone of diagnostic AI tools, risk stratification models, and more.

**Anomaly Detection:** The use of AI to identify patterns in data that deviate from the norm, often used in fraud detection or patient deterioration alerts.

**Artificial Intelligence (AI):** A field of computer science focused on creating systems that simulate human intelligence to perform tasks such as diagnosis, triage, imaging analysis, and administrative automation.

**Augmented Intelligence:** A collaborative model where AI supports and enhances human decision-making rather than replacing it—widely promoted in healthcare ethics.

**Bayesian Network:** A probabilistic model used for reasoning and decision-making under uncertainty; used in some AI diagnostic systems.

**Benchmark Dataset:** A standard dataset used to compare the performance of AI models—e.g., MIMIC-III for critical care AI research.

**Bias (Algorithmic):** Systemic deviation in data or algorithms leading to skewed outcomes; may result in disparities across gender, race, age, or condition.

**Big Data:** Extremely large datasets that are analyzed computationally to reveal patterns and trends, often used to train healthcare AI models.

**Black Box:** An AI model whose internal logic and decision-making are not transparent or understandable by humans—a challenge in clinical AI.

**Chatbot:** An AI tool that interacts with users via natural language, used in patient intake, symptom checking, or mental health support.

**Clinical Decision Support (CDS):** AI-powered tools that provide healthcare professionals with knowledge and patient-specific information to enhance decisions.

**Clinical Decision Support System (CDSS):** Software that uses AI to assist healthcare providers in decision-making by offering evidence-based recommendations or alerts.

**Cognitive Computing:** An advanced form of AI mimicking human thought processes in complex problem-solving—used in diagnostics and personalized care.

**Computer Vision:** A field of AI that allows machines to interpret and make decisions based on visual input like X-rays or MRIs.

**Data Drift:** A change in the statistical properties of input data over time that can degrade AI performance if not monitored.

**Data Governance:** Framework for managing data availability, usability, integrity, and security—critical in healthcare AI use.

**Data Lake:** A centralized repository for storing large volumes of structured and unstructured healthcare data used in training AI models.

**Deep Learning:** A subset of machine learning using multi-layered neural networks for highly complex pattern recognition (e.g., tumor detection in imaging).

**Digital Twin:** A virtual replica of a physical system (e.g., a patient or organ) used to simulate outcomes, interventions, or disease progression.

**EHR Integration:** The process of embedding AI tools directly within electronic health records systems for seamless clinical use.

**Edge AI:** AI computation performed locally (e.g., on a wearable device) rather than in the cloud, allowing real-time feedback with privacy benefits.

**Electronic Health Record (EHR):** A digital version of a patient's medical history. Often used as a data source for training or deploying AI models.

**Explainable AI (XAI):** Methods that make AI systems interpretable to humans. Crucial for accountability in healthcare decisions (e.g., why a diagnosis was suggested).

**False Positive / False Negative:** Incorrect outputs from an AI model—false positives flag a condition that isn't there; false negatives miss a condition that is.

**Feature Extraction:** The process of transforming raw data into measurable and meaningful signals (features) for AI model training.

**Federated Learning:** A machine learning approach that trains algorithms across decentralized data sources without sharing patient data, helping preserve privacy.

**Generative AI:** AI models (like GPT) capable of creating new content such as text, images, or simulations—used in healthcare education, documentation, and training.

**Ground Truth:** The real, verified data used to train and test AI models for accuracy.

**GPU (Graphics Processing Unit):** Specialized hardware often used in healthcare AI to process large amounts of data, especially in deep learning tasks.

**Health Informatics:** The science of collecting, analyzing, and using health data and AI to improve patient outcomes.

**HIPAA Compliance:** Ensuring AI systems handling personal health data adhere to U.S. privacy and security regulations.

**Human-in-the-Loop (HITL):** A model where humans validate or influence AI decisions, used in clinical workflows to ensure safety.

**Imputation:** The technique of filling in missing healthcare data points, often used in preparing training datasets.

**Inference Engine:** The part of an AI system that applies rules or models to input data to generate conclusions or predictions.

**Input Data:** Data used to feed an AI system—must be clean, accurate, and ethically sourced in healthcare.

**Interoperability:** The ability of AI systems to exchange and use data across different healthcare systems.

**Knowledge Graph:** A structured representation of relationships between medical concepts (e.g., symptoms, diseases, drugs) to improve AI reasoning.

**Labeling (Data):** The process of tagging data with meaningful labels (e.g., “COVID positive”) for supervised learning.

**Logistic Regression:** A statistical method commonly used in clinical prediction models and early-stage machine learning.

**Machine Learning (ML):** A subset of AI that enables systems to learn from data and improve performance without explicit programming.

**Model Drift:** When an AI model’s performance decreases over time due to changing data patterns—must be monitored in clinical AI.

**Model Interpretability:** The ability to understand how an AI model arrives at its decisions—important for clinician trust.

**Model Training:** The phase where an AI system learns patterns from labeled data before being deployed in clinical settings.

**Multimodal AI:** AI that integrates multiple data types (e.g., text, image, and lab results) to make more accurate decisions.

**Natural Language Processing (NLP):** A branch of AI that enables machines to understand and generate human language, crucial for analyzing medical records and patient communication.

**Neural Network:** A type of algorithm modeled after the human brain, used in deep learning for pattern recognition.

**Noise (Data):** Irrelevant or incorrect data that may reduce AI model accuracy.

**Ontology:** A formal naming and definition system for the relationships between healthcare concepts, used in structuring AI knowledge.

**Overfitting:** When an AI model performs well on training data but poorly on new data—commonly mitigated by regularization and validation.

**Personalized Medicine:** Tailoring medical treatment to individual patient characteristics using AI to analyze genetic, clinical, and lifestyle data.

**PHI (Protected Health Information):** Any health data that can identify a patient—must be handled with care under HIPAA and AI privacy guidelines.

**Precision Health:** A proactive approach using AI to prevent disease and optimize care based on predictive risk models.

**Predictive Analytics:** AI-driven techniques that forecast future outcomes based on historical data—e.g., predicting patient deterioration.

**Prompt Engineering:** The practice of crafting effective inputs for generative AI models to yield desired outputs, especially in documentation or education tools.

**Quality Assurance (QA):** Measures taken to ensure AI systems in healthcare function reliably, including validation, peer review, and continuous monitoring.

**Regression Model:** An AI model that predicts continuous outcomes (e.g., length of hospital stay or readmission probability).

**Reinforcement Learning:** An AI method where agents learn by receiving rewards or penalties for their actions—less common but emerging in personalized care.

**Responsible AI:** The practice of building and using AI systems in a way that is fair, accountable, and aligned with ethical standards.

**Risk Stratification:** AI use in categorizing patients by their risk levels to guide interventions or care management.

**Robotic Process Automation (RPA):** Software that mimics human actions to automate repetitive tasks—used in billing, scheduling, or claims processing.

**Scalability:** The ability of an AI solution to handle increasing data or expand across healthcare settings without loss of performance.

**Semi-Supervised Learning:** Combines a small amount of labeled data with a larger volume of unlabeled data, improving training efficiency.

**Structured Data:** Data organized in a defined format, such as lab results or vitals—easier for AI models to process.

**Supervised Learning:** An AI training method using labeled datasets—common in diagnostics and triage tools.

**Synthetic Data:** Artificially generated data used to augment datasets while preserving privacy and reducing bias risks.

**Training Data:** Data used to teach an AI system to recognize patterns or make decisions.

**Transfer Learning:** Applying knowledge from one trained model to a new but related problem—useful in low-data environments like rare diseases.

**Transparency:** The ability to understand how an AI system functions and makes decisions—a core tenet of ethical AI.

**Trust Calibration:** Ensuring healthcare users neither under-trust nor over-trust AI tools, promoting appropriate use through design and education.

**Unstructured Data:** Data not in a predefined format—e.g., clinical notes, radiology reports. Requires NLP or other techniques for analysis.

**Unsupervised Learning:** AI systems trained without labeled outputs—used to discover hidden patterns or clusters (e.g., patient phenotypes).

**Usability Testing:** Evaluating how easily clinicians and staff can use an AI tool—important for adoption and safety.

**Use Case:** A specific scenario where AI is applied to solve a healthcare problem or improve operations.

**Validation (AI Model):** The process of testing an AI model to ensure it performs as intended with new, unseen data.

**Validation Dataset:** A dataset used to tune the model during training, helping prevent overfitting.

**Value Analysis (in AI):** The process of evaluating the cost, benefit, risk, and clinical effectiveness of implementing an AI tool in healthcare.

**Value-Based Care:** A model of care focused on outcomes rather than services—AI is increasingly used to support this transition.

**Wearable Technology:** Devices like smartwatches or biosensors that collect real-time health data used by AI for monitoring and alerts.

**Workflow Integration:** The process of embedding AI into clinical or operational processes to improve efficiency without disrupting care delivery.

**Workflow Optimization:** The use of AI to improve operational efficiency, reduce bottlenecks, or enhance care coordination.

**XAI (Explainable Artificial Intelligence):** See “Explainable AI.” Emphasized in clinical settings for trust and compliance.

**Zero-shot Learning:** A method where AI applies knowledge to new problems without explicit training on those categories—emerging in diagnostic AI.