ORACLE

Developing Trustworthy AI in Healthcare

Oracle Health and Life Sciences' approach to transparent and accountable AI

June, 2026

Public

ORACLE

Purpose statement

This document describes Oracle Health and Life Sciences' approach to trustworthy AI, including the governance framework, guiding principles, and operational controls used to support transparency, accountability, reliability, privacy, security, fairness, and human oversight throughout the AI lifecycle.

Disclaimer

This document in any form, software or printed matter, contains proprietary information that is the exclusive property of Oracle. Your access to and use of this confidential material is subject to the terms and conditions of your Oracle software license and service agreement, which has been executed and with which you agree to comply. This document and information contained herein may not be disclosed, copied, reproduced or distributed to anyone outside Oracle without prior written consent of Oracle. This document is not part of your license agreement nor can it be incorporated into any contractual agreement with Oracle or its subsidiaries or affiliates.

This document is for informational purposes only and is intended solely to assist you in planning for the implementation and upgrade of the product features described. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, timing, and pricing of any features or functionality described in this document remains at the sole discretion of Oracle. Due to the nature of the product architecture, it may not be possible to safely include all features described in this document without risking significant destabilization of the code.

This document may include some forward-looking content for illustrative purposes only. Some products and features discussed are indicative of the products and features of a prospective future launch in the United States only or elsewhere. Not all products and features discussed are currently offered for sale in the United States or elsewhere. Products and features of the actual offering may differ from those discussed in this document and may vary from country to country. Any timelines contained in this document are indicative only. Timelines and product features may depend on regulatory approvals or certification for individual products or features in the applicable country or region.

ORACLE

Table of contents

 3580184312_Developing Trustworthy AI in Healthcare Business Brief_v1 | June 2026

Introduction

When developed and deployed responsibly, artificial intelligence (AI) has the potential to fundamentally transform the delivery and management of healthcare. AI can help clinicians synthesize complex patient information more efficiently, support more informed and timely decision-making, reduce administrative burden, and enable healthcare organizations to operate more effectively at scale. These capabilities can not only improve the quality and consistency of care but also create opportunities to reduce inefficiencies and overall healthcare costs.

Realizing this potential requires AI to be deployed with rigorous oversight and governance, particularly given its potential impact on patient safety and direct influence on the trust of patients, clinicians, payers, and regulators. Inaccurate or inapplicable outputs can lead to misdiagnosis, inappropriate treatment and coding, and delays in care, especially when underlying data is incomplete or unrepresentative. There are also significant concerns around patient privacy, data security, and the potential for unauthorized use of sensitive health information.

For these reasons, healthcare AI is subject to strict regulatory and compliance requirements globally. These may include legal and regulatory frameworks such as the Health Insurance Portability and Accountability Act (HIPAA) in the United States (US) and the AI Act and the General Data Protection Regulation (GDPR) in the European Union (EU), as well as a myriad of requirements across individual US states and federal regulators. Oversight is provided by regulatory authorities including the UK Medicines and Healthcare products Regulatory Agency (MHRA) and the EU Member States jurisdiction.

To support compliance with these requirements, Oracle integrates AI governance and risk management into its broader software development lifecycle, including review, testing, validation, and monitoring processes throughout the AI lifecycle. In addition, Oracle Health and Life Sciences has taken supplementary steps to support the structured review of AI features.

Oracle Health and Life Sciences has established a formal governance framework which provides a structured review process for all AI features prior to their release. The framework is grounded in a set of core principles that guide how AI systems are designed, deployed, governed, and continuously monitored throughout their lifecycle with a keen eye on the unique factors for AI in a healthcare setting. Together, these principles establish the foundation for how Oracle Health and Life Sciences operationalizes trustworthy AI across clinical, operational, and administrative environments.

Our governance framework requires a comprehensive evaluation of each AI model's design, performance, and compliance with applicable standards, such as fairness, explainability, privacy, and security. It includes reviews of model purpose, intended users, data provenance, grounding strategy, safety risks, reliability evidence, human oversight design, monitoring plans, and feedback mechanisms. Furthermore, this framework has been directly integrated into Oracle Health's broader software development lifecycle and quality management system and has achieved ISO/IEC 42001 AI Management System (AIMS) certification.

By combining technical safeguards with formal governance processes, Oracle Health and Life Sciences has established an approach designed to support adherence to regulatory requirements, high standards for quality assurance, and broader expectations for transparency, oversight, and accountability. This layered approach is essential for maintaining trust in environments where data sensitivity and regulatory scrutiny are high.

Principles for trustworthy AI

Oracle AI Review for Health is guided by a set of core principles that define the key requirements for responsible and trustworthy AI in design, development, and oversight. These principles are human-centered design, transparency, performance and reliability, privacy and security, fairness and integrity, and accountability.

Representative examples are included below to illustrate how certain principles are applied in practice, particularly where real-world context helps clarify how AI operates within clinical and operational workflows.

Principle 1: Human-centered design

AI augments decision-making, but humans remain accountable and in control. Oracle Health and Life Sciences' AI systems are designed to support human-in-the-loop workflows, so that clinicians, administrators, and financial teams review AI outputs before final decisions are made and before actions are taken. This approach reinforces the core idea that AI serves as a support tool rather than a replacement for professional judgment.

In practice, this approach means AI-generated outputs are typically presented as draft recommendations within existing workflows with users reviewing and validating results before acting. This approach helps maintain clear accountability and supports decisions reflect both machine-generated insights and human expertise.

For example, in clinical documentation workflows, AI may generate a draft discharge summary of a patient encounter, but a clinician must review and approve the content before it becomes part of the official medical record. This human-plus-AI approach can help support quality outcomes.

Additionally, relevant subject matter experts (SMEs)—most specifically clinicians—are involved in the actual design process for our AI systems, which is explicitly architected with multiple checkpoints for input from SMEs and potential end users. Clinician feedback is key to developing our evaluation framework for each AI system, and we ask clinicians to annotate our evaluation/validation sets whenever there is any room for clinical interpretation in outputs.

Principle 2: Transparency and explainability

AI outputs are designed to be transparent and traceable, enabling users to understand how the AI results are generated and verify their accuracy. Rather than presenting conclusions as opaque outputs, Oracle Health and Life Sciences' AI systems are designed to provide visibility into the data sources, contributing factors, and logic underlying each result, making the outputs explainable for users. For example, clinicians can trace elements of an AI-generated summary back to clinical notes, lab results, well established clinical guidelines, or other customer-approved data sources, enabling rapid verification and human oversight, and thus reinforcing confidence in the system.

This "show your work" approach is particularly important in healthcare, where clinicians and administrators must be able to evaluate and trust the information they use. While AI is inherently probabilistic and, therefore, will never be 100% perfect, transparency provides the user with the information needed to quickly assess whether an output is complete, relevant, and appropriate for the context in which it is being used, and feel confident when/if they disagree with an output.

We also employ transparency using model cards for AI systems and models. Model cards are a transparency artifact concept used for supplying standardized documentation describing all the information a user needs to know about an AI tool to make an informed use of it. This documentation includes details such as the intended and non-intended uses, training and evaluation data sets, performance measures, and ongoing maintenance procedures. These model cards are made available for Oracle Health and Life Sciences' AI systems and models both directly within the electronic health record (EHR) system for review by clinicians at the point of care, as well as in product documentation.

Principle 3: Performance and reliability

AI systems and models are continuously evaluated, validated, and governed throughout their lifecycle to support accuracy, reliability, and alignment with clinical, financial, and operational standards. This evaluation and governance framework includes both pre-deployment validation and ongoing monitoring to assess whether models continue to perform as expected in real-world environments. Beyond performance, model governance also helps support AI models remaining effective, fit for purpose, and useful in the workflows they are intended to support.

Model governance is enforced through structured processes, such as our framework, which establishes the criteria for evaluating model performance, fairness, reliability, and compliance prior to release. These reviews include evidence from quantitative evaluations, bias and fairness assessments (including subgroup analyses where applicable) robustness assessments, and reliability assessments.

After deployment, models are monitored through defined performance metrics and feedback mechanisms with thresholds that trigger review, retraining, or corrective action when necessary. For example, a revenue cycle model may be monitored for coding accuracy, documentation completeness, and compliance risks to help maintain alignment with regulatory and operational requirements over time. This monitoring helps detect output drift, changes in model behavior, unexpected shifts in usage patterns, or emerging quality concerns. These signals can be routed to human reviewers for triage, prioritization, and corrective action creating a feedback loop between production performance, expert review, and continuous model improvement.

Critically, these safety and efficacy evaluations are performed specific to the intended use and/or user of the model. Every AI system and model is unique, and while a common framework for review and governance is important, the evaluations must also be appropriately tailored to the individual model for quality and effectiveness.

Principle 4: Security and privacy

Data is protected through privacy-by-design principles, meaning privacy safeguards are built into AI systems from the outset, along with secure architecture and operational controls. This includes measures such as data minimization, de-identification, encryption, and strict access controls to protect sensitive health information throughout its lifecycle.

Oracle aligns its data handling practices with global regulatory frameworks such as HIPAA in the US, GDPR in the EU and the UK GDPR in the UK. These frameworks guide how data is collected, used, stored, and shared to support patient privacy protections and organizations' regulatory compliance efforts. Oracle's distributed cloud capabilities also support data sovereignty requirements by enabling organizations to store and process data within required geographic or regulatory boundaries.

By embedding privacy and security into system design rather than treating them as afterthoughts, Oracle helps reduce risk and build trust with patients, providers, and regulators.

Principle 5: Fairness and integrity

AI models are designed, deployed, and monitored to support accurate outputs across both patient populations and administrative processes. This approach includes evaluating model performance across demographic groups and assessing outputs for potential disparities in care or operational outcomes, as well as introducing post-deployment controls where appropriate. These activities are incorporated into our governance framework, so fairness is evaluated as part of release readiness rather than treated as a separate or optional activity.

In addition to clinical fairness, this principle also encompasses administrative integrity. AI systems are evaluated to identify inconsistencies or risks in areas such as coding accuracy, billing practices, and compliance with payer requirements as appropriate and necessary for the given AI system or model. These assessments help limit the risk of issues such as improper payments or systematic errors that could impact both financial performance and regulatory compliance.

Principle 6: Accountability

Clear ownership is established for AI systems, with defined roles and responsibilities intended to support accountability for AI systems across the lifecycle. Accountability applies across the full lifecycle of an AI system, from development and validation to deployment and ongoing use.

In practice, this accountability framework includes assigning responsibility for model development, oversight, and performance monitoring, as well as, when appropriate, implementing role-based access controls that determine who can review, approve, and act on AI-generated outputs within a healthcare organization. These controls are intended to support decision-making by authorized individuals and help maintain clear accountability for those decisions.

By maintaining clear lines of accountability, Oracle supports responsible use of AI and helps organizations govern and manage AI-enabled workflows effectively.

ORACLE

Putting the principles into practice

The principles described above are implemented through a combination of technical controls, operational governance processes, and secure infrastructure. The following examples illustrate how Oracle Health and Life Sciences translates its trustworthy AI principles into day-to-day practices.

Multimodal data and model design

AI systems are designed to integrate diverse sources and healthcare data types—including structured data such as lab results and medications, unstructured clinical notes, operational inputs, and administrative requirements—to generate comprehensive and contextually relevant outputs. This multimodal approach reflects the complexity of real-world healthcare environments and enables AI systems to capture clinical and operational context.

Oracle's AI approach enables organizations to incorporate localized operational, administrative, and clinical information, such as internal treatment guidelines, payer authorization requirements, formulary restrictions, and workflow protocols, so outputs align with the organization's specific practices and environment. Where applicable, AI outputs are grounded in governed knowledge sources, including Oracle Health's semantic knowledge graph, including validated ontology, terminology, and relationships data. By combining these different data sources and constraints, AI systems can support informed, operationally relevant, and accurate recommendations. This approach also helps maintain AI outputs as grounded in validated enterprise and clinical context rather than relying solely on generalized model responses.

In practice, different types of AI models may be used depending on the task, including language models for summarization, retrieval-based systems for accessing relevant enterprise data, and reasoning models for organizing and synthesizing information. Where applicable, enterprise data sources and reference content are refreshed and updated through defined governance and integration processes to help keep operational and administrative inputs remain current.

Data governance and responsible data use

Data used for AI development and operation is governed through defined controls on access, management, and use, including role-based permissions, data handling requirements, and responsible data practices. These controls are intended to support appropriate, secure, and compliant data use in alignment with regulatory, contractual, and organizational requirements.

Depending on the product, Oracle may use a combination of de-identified healthcare data, permitted public datasets, licensed medical references, synthetic datasets, and/or customer-authorized enterprise data for model training, evaluation, and operational workflows. Publicly available and reference sources may include peer-reviewed biomedical literature, clinical practice guidelines from recognized medical organizations, public health resources, standardized medical terminologies such as SNOMED CT (Systematized Nomenclature of Medicine – Clinical Terms) and ICD (International Classification of Diseases) coding systems, and publicly available Centers for Medicare & Medicaid Services' (CMS) coding and reimbursement guidance relevant to revenue cycle workflows.[1]

By default, data used for training and evaluation is de-identified in accordance with HIPAA standards, with alignment to global frameworks such as GDPR where applicable. Reference datasets, coding guidance, and operational rules are maintained and updated through defined governance, integration, and content management processes to support the ongoing use of current and relevant information by AI systems. All data access is controlled through role-based permissions, and data use is subject to safeguards designed to prevent misuse or unauthorized disclosure.

Cybersecurity and secure AI operations

AI systems are deployed within secure environments that protect against unauthorized access, threats, and vulnerabilities, while helping maintain system integrity across the AI lifecycle. This security framework includes

[1] Examples may include peer-reviewed journals and databases such as PubMed/MEDLINE, publicly available guidance from organizations such as the CDC and NIH, specialty society clinical guidelines, and publicly available CMS coding and reimbursement resources.

secure cloud infrastructure, authentication and authorization controls, continuous monitoring, and threat detection capabilities.

Cybersecurity focuses on protecting the systems that process and generate AI outputs, complementing data privacy protections by addressing risks such as unauthorized system access, malicious activity, and operational disruptions.

A key component of this approach is minimizing unnecessary data movement. Rather than transferring sensitive data to external systems, Oracle prioritizes bringing AI capabilities to data within governed enterprise environments, helping minimize exposure risk and maintain control over sensitive information. This architecture supports both operational resilience and regulatory compliance while reinforcing customer trust in AI-enabled healthcare systems.

Together, these principles reflect Oracle Health and Life Sciences' commitment to developing and deploying AI in a manner that is transparent, and accountable. Combined with our governance framework, they establish a comprehensive approach for evaluating, governing, and continuously improving AI systems across their lifecycle. This approach reflects Oracle Health and Life Sciences' leadership in advancing transparent and accountable AI practices in healthcare, with an emphasis on balancing innovation with strong safeguards, oversight, and trust.

Conclusion

Oracle Health and Life Sciences has developed a comprehensive approach to AI that integrates governance, transparency, and accountability across the full lifecycle of its models. By embedding AI into workflows, enforcing strong data governance practices, and enabling traceability and human oversight, Oracle Health and Life Sciences supports informed decision-making and responsible AI use among clinicians, administrative professionals, and patients.

As AI continues to play a larger role in healthcare, maintaining trust will be essential for adoption and impact. Oracle's approach demonstrates that trustworthy AI requires more than technical performance alone. It depends on coordinated governance, transparent operations, continuous oversight, responsible data practices, and the consistent application of core AI principles across the full AI lifecycle.