

# AI Ethics and Security in Healthcare: Frameworks for Responsible AI Implementation



From streamlining patient care to enhancing clinical decision-making, healthcare AI technologies hold tremendous potential. However, with this potential comes ethical and security risks that must be managed at each step of the adoption journey.

This guide shares selected information about the World Health Organization’s (WHO) AI ethics framework and the Open Worldwide Application Security Project (OWASP) AI security guidelines. From safeguarding patient data to preventing cybersecurity threats, this reference is designed to help healthcare organizations prioritize responsible, compliant and secure AI practices.

*Aidoc makes no ownership claim over the material presented in this document. These streamlined guides have been adapted from publicly available resources to serve as a reference supporting best practices for responsible and secure AI adoption.*

## The WHO AI Ethics Framework

The table<sup>1</sup> below translates the **WHO AI ethics framework** into practical, actionable steps tailored to each phase of the AI lifecycle. It outlines clear responsibilities for AI developers and corresponding actions for governments, ensuring alignment with ethical principles and regulatory standards.

AI STAGE CONSIDERATIONS	DEVELOPER ACTIONS	GOVERNMENT ACTIONS
<b>Development Phase</b>		
 Bias	 Privacy	<ul style="list-style-type: none"> <li>• Provide certification/training for programmers</li> <li>• Have and enforce strong data protection laws</li> </ul>
 Labor Concerns	 Carbon and Water Footprints	<ul style="list-style-type: none"> <li>• Issue target product profiles</li> <li>• Mandate outcomes (predictability, interpretability, corrigibility, safety and cybersecurity)</li> </ul>
 False Information or Misinformation	 Safety and Cybersecurity	<ul style="list-style-type: none"> <li>• Ensure refreshed, updated and context-appropriate training data</li> <li>• Ensure transparency of training data</li> <li>• Introduce pre-certification programs to identify and avoid ethical risks</li> </ul>
 Epistemic Authority of Humans	 Exclusive Control of LLMs	<ul style="list-style-type: none"> <li>• Offer fair wages and support to data workers</li> <li>• Conduct audits during early AI development</li> <li>• Involve diverse stakeholders in design</li> <li>• Require developers to address carbon and water footprints</li> <li>• Design for accuracy and predictability</li> <li>• Require developers to label AI-generated content for users</li> </ul>

# AI Ethics and Security in Healthcare: Frameworks for Responsible AI Implementation











AI STAGE CONSIDERATIONS	DEVELOPER ACTIONS	GOVERNMENT ACTIONS
<b>Development Phase</b>		
	<ul style="list-style-type: none"> <li>Design for values based on consensus principles and ethical norms</li> <li>Design to improve energy efficiency of models</li> </ul>	<ul style="list-style-type: none"> <li>Encourage or require early-stage registration of algorithms</li> <li>Invest in or provide public or not-for-profit infrastructure</li> <li>Promote open-source LLMs</li> </ul>

AI STAGE CONSIDERATIONS	DEVELOPER ACTIONS	GOVERNMENT ACTIONS
<b>Provision Phase</b>		
 System-Wide Bias	 False Information or Misinformation	<ul style="list-style-type: none"> <li>Assign regulatory agency to assess and approve LLMs for health</li> <li>Require transparency, including source code and data inputs</li> </ul>
 Manipulation	 Privacy	<ul style="list-style-type: none"> <li>Enforce data protection laws for user-inputted data</li> <li>Mandate ethical and human rights standards, irrespective of risk or benefit</li> </ul>
 Automation Bias		<ul style="list-style-type: none"> <li>Enact laws requiring impact assessments, audited by third parties and disclosed publicly</li> <li>Prohibit non-trial experimental use; explore regulatory sandboxes for controlled testing</li> <li>Require developers to label AI-generated content for users</li> <li>Apply consumer protection laws to prevent negative impacts on end-users and patients</li> </ul>

# AI Ethics and Security in Healthcare: Frameworks for Responsible AI Implementation






AI STAGE CONSIDERATIONS	DEVELOPER ACTIONS	GOVERNMENT ACTIONS
<b>Deployment Phase</b>		
 Inaccurate or False Responses	 Bias	<ul style="list-style-type: none"> <li>Avoid using LLMs in inappropriate settings</li> <li>Mandate independent post-release audits and impact assessments for LMM deployment</li> </ul>
 Privacy	 Accessibility and Affordability	<ul style="list-style-type: none"> <li>Communicate known risks, errors and harms with clear warnings and measures</li> <li>Hold developers or providers responsible for false or toxic information</li> <li>Enforce operational disclosures, including technical documentation</li> </ul>
 Labor and Employment	 Automation Bias	<ul style="list-style-type: none"> <li>Train healthcare workers on LLM decision-making, avoiding bias, patient engagement and cybersecurity risks</li> </ul>
 Quality of Clinician-Patient Interaction	 Skills Degradation	<ul style="list-style-type: none"> <li>Facilitate public participation through human oversight colleges to ensure appropriate use</li> <li>Engage the public to understand data sharing, assess social/cultural acceptability, improve AI literacy and gauge acceptable LLM uses</li> <li>Use procurement authority to encourage transparency and responsible practices by value chain actors</li> </ul>

# AI Ethics and Security in Healthcare: Frameworks for Responsible AI Implementation





## WHO AI Ethics Framework: Benefits and Risks of Large Language Models (LLMs) in Healthcare

Outlined below are the benefits and risks of using LLMs in healthcare as noted in the WHO AI ethics framework. Covering applications like diagnosis, patient support, administration, education and research it helps clinicians, administrators, developers and policymakers make informed decisions about leveraging LLMs responsibly and aligning their adoption with ethical and secure practices.







USE CASE	POTENTIAL BENEFITS	POTENTIAL RISK
 Diagnosis and Clinical Care	<ul style="list-style-type: none"> <li>Assist in managing complex cases and routine diagnoses</li> </ul>	<ul style="list-style-type: none"> <li>Inaccurate, incomplete or false responses</li> </ul>
	<ul style="list-style-type: none"> <li>Reduce communication workload (“keyboard liberation”)</li> </ul>	<ul style="list-style-type: none"> <li>Poor quality training data and bias</li> </ul>
	<ul style="list-style-type: none"> <li>Provide novel insights from unstructured health data</li> </ul>	<ul style="list-style-type: none"> <li>Automation bias</li> </ul>
		<ul style="list-style-type: none"> <li>Skill degradation in healthcare professionals</li> <li>Informed consent challenges</li> </ul>
 Patient-Guided Use	<ul style="list-style-type: none"> <li>Improve understanding of medical conditions (patients or caregivers)</li> </ul>	<ul style="list-style-type: none"> <li>Inaccurate or misleading statements</li> </ul>
	<ul style="list-style-type: none"> <li>Use a virtual health assistant</li> </ul>	<ul style="list-style-type: none"> <li>Privacy concerns</li> </ul>
	<ul style="list-style-type: none"> <li>Support clinical trial enrollment</li> </ul>	<ul style="list-style-type: none"> <li>Reduced clinician-patient interactions</li> <li>Epistemic injustice</li> </ul>
		<ul style="list-style-type: none"> <li>Risk of care delivery outside established health systems</li> </ul>
 Clerical and Administrative Tasks	<ul style="list-style-type: none"> <li>Streamline paperwork and clinical documentation</li> </ul>	<ul style="list-style-type: none"> <li>Potential inaccuracies or errors</li> </ul>
	<ul style="list-style-type: none"> <li>Translate languages</li> </ul>	<ul style="list-style-type: none"> <li>Inconsistent responses to varying prompts</li> </ul>
	<ul style="list-style-type: none"> <li>Automate electronic health record updates</li> </ul>	
	<ul style="list-style-type: none"> <li>Draft clinical notes post-visit</li> </ul>	

# AI Ethics and Security in Healthcare: Frameworks for Responsible AI Implementation



USE CASE	POTENTIAL BENEFITS	POTENTIAL RISK
 Medical and Nursing Education	<ul style="list-style-type: none"> <li>Tailor adaptive educational texts to students</li> </ul>	<ul style="list-style-type: none"> <li>Automation bias in learning</li> </ul>
	<ul style="list-style-type: none"> <li>Incorporate simulated conversations for practice</li> </ul>	<ul style="list-style-type: none"> <li>Errors or false information affecting education quality</li> </ul>
	<ul style="list-style-type: none"> <li>Provide reasoned responses for learning</li> </ul>	<ul style="list-style-type: none"> <li>New digital skill requirements for educators and students</li> </ul>
 Scientific Research and Drug Development	<ul style="list-style-type: none"> <li>Analyze and generate insights from research data</li> </ul>	<ul style="list-style-type: none"> <li>Accountability for algorithm-generated content</li> </ul>
	<ul style="list-style-type: none"> <li>Draft scientific articles</li> </ul>	<ul style="list-style-type: none"> <li>Bias in data favoring high-income countries</li> </ul>
	<ul style="list-style-type: none"> <li>Proofread and summarize</li> </ul>	<ul style="list-style-type: none"> <li>Generating false or non-existent references</li> </ul>
	<ul style="list-style-type: none"> <li>Aid in de novo drug design</li> </ul>	<ul style="list-style-type: none"> <li>Undermining peer review and scientific rigor</li> </ul>

## Systemic Risks of LLMs in Healthcare

TYPE OF RISK	DESCRIPTION
 <b>Overestimation of Benefits</b>	Over-reliance on LLMs may lead to “technological solutionism,” downplaying safety, efficacy and utility challenges.
 <b>Accessibility and Affordability</b>	Lack of equitable access due to the digital divide or high subscription fees could widen disparities in care.
 <b>System-Wide Biases</b>	Larger datasets may encode biases that get perpetuated throughout healthcare systems.
 <b>Labor Impact</b>	Job loss and the need for workforce retraining may arise, alongside poor working conditions for data annotators.
 <b>Dependence on Ill-Suited LLMs</b>	Low-maintenance or regionally biased LLMs can erode trust and privacy in healthcare systems.
 <b>Cybersecurity Risks</b>	Malicious attacks or hacking could compromise the safety and trust in LLM-dependent healthcare systems.

# AI Ethics and Security in Healthcare: Frameworks for Responsible AI Implementation



## OWASP LLM AI Cybersecurity and Governance Checklist: A Quick Reference Guide

The OWASP Top 10 Checklist helps leaders in tech, cybersecurity and compliance develop secure AI strategies while mitigating risks. It offers actionable insights aligned with global standards like GDPR and the EU AI Act, ensuring safe and effective AI adoption. Below are select sections from the checklist that complement the WHO AI ethics framework.

Explore the [full OWASP AI resource](#) for detailed strategies and best practices.<sup>2</sup>

### Adversarial Risk

- Competitor Analysis:** Scrutinize how competitors invest in AI to evaluate market impacts and opportunities.
- Defensive Controls:** Assess and update current controls, such as voice-based password resets, to counter attacks enhanced by generative AI (GenAI).
- Incident Response:** Revise playbooks and plans to address machine learning (ML) and AI-specific threats and attacks.

### Threat Modeling

- Systematic Risk Identification:** Use threat modeling to examine vulnerabilities and processes for AI systems.
- GenAI-Accelerated Threats:** Anticipate LLM-assisted phishing and hyper-personalized attacks.
- Customer Protection:** Mitigate risks from spoofing or malicious GenAI-generated content targeting clients.
- Insider Threat Mitigation:** Implement safeguards against misuse by authorized users.
- Content Filtering:** Automate mechanisms to prevent harmful outputs from AI systems.

### AI Asset Inventory

- Comprehensive Cataloging:** Identify all AI tools, services and data sources, tagging them for sensitivity and ownership.
- Software Bill of Materials (SBOM):** Include AI components and dependencies for detailed tracking and management.
- Risk Assessments:** Conduct pen testing and red teaming to evaluate attack surfaces.
- Onboarding Processes:** Establish protocols for adopting AI solutions and ensure alignment with IT expertise.

# AI Ethics and Security in Healthcare: Frameworks for Responsible AI Implementation



## AI Security and Privacy Training

- **Employee Engagement:** Address employee concerns transparently about LLM initiatives.
- **Ethical and Legal Training:** Educate users on AI-related ethics, warranties, licenses and copyright issues.
- **GenAI Threat Awareness:** Update security training to address risks like voice cloning, image cloning and spear phishing.
- **DevOps and Cybersecurity:** Train teams on safe deployment and security assurances for AI tools.



## Establish Business Cases

- **Strategic AI Adoption:** Develop strong business cases balancing risk, benefits and ROI for AI initiatives.
- **Use Case Examples:** Focus on customer experience, operational efficiency, market research, innovation and document management.



## Governance

- **Accountability Framework:** Create an AI RACI chart to define responsibilities and ensure transparency.
- **Data Management:** Implement policies for secure data classification and usage restrictions.
- **AI Policies:** Develop acceptable use matrices and establish guidelines for generative AI tools.
- **Source Management:** Document the sources and processes governing GenAI data.



## Legal

- **Liability and Warranties:** Define clear terms in end-user license agreements to address AI-generated risks.
- **Intellectual Property:** Safeguard proprietary content and address risks from AI-assisted code generation.
- **Regulatory Alignment:** Ensure AI tools comply with laws regarding bias, plagiarism and data privacy.
- **Risk Mitigation:** Establish guardrails for indemnification provisions and evaluate insurance coverage adequacy.

# AI Ethics and Security in Healthcare: Frameworks for Responsible AI Implementation



## Regulatory

- Jurisdictional Compliance:** Identify and adhere to AI-specific regulations, such as the EU AI Act and GDPR.
- Employment AI Tools:** Evaluate tools used for hiring to ensure fairness, bias mitigation and data privacy.
- Vendor Compliance:** Confirm third-party adherence to AI regulations and best practices.

## Using or Implementing LLM Solutions

- Trust Boundary Security:** Threat model LLM components and secure integrations.
- Access and Data Controls:** Enforce least privilege access and classify data based on sensitivity.
- Pipeline Governance:** Ensure rigorous control over training data and algorithm security.
- Incident Response:** Update response playbooks to include LLM-specific attacks and vulnerabilities.
- Audit and Monitoring:** Establish processes for automating, logging and auditing workflows.

## Testing, Evaluation, Verification and Validation (TEVV)

- Lifecycle Continuity:** Implement ongoing TEVV processes to maintain AI system reliability and security.
- Executive Oversight:** Deliver regular updates on AI model performance, risks and robustness.
- Proactive Adjustments:** Recalibrate and monitor AI systems periodically to adapt to evolving risks.

## Explore More AI Governance Resources

### On-Demand Webinars

“Regulating the Future: A Deep Dive into Healthcare AI Governance” features healthcare and legal experts from Deloitte Consulting, American College of Cardiology and Epstein Becker Green explaining potential approaches to governance and essential considerations.

[WATCH WEBINAR](#)

### Resource Guides

The “5 Essential Areas of Alignment for Governing Clinical AI Partnerships” identifies key collaboration points with AI developers, while the “Clinical AI Readiness Assessment” helps evaluate healthy system readiness for scalable and secure AI, highlighting strengths and areas for improvement.

[VIEW RESOURCES](#)

### AI Learning Center

The path from “what if” to “what’s next” requires careful planning. It’s why we’ve curated a Governance section within our AI Learning Center to provide insights on security considerations and regulatory updates to guide your journey.

[EXPLORE MORE](#)



©aidoc | [info@aidoc.com](mailto:info@aidoc.com) | For safety information on Aidoc’s products, please visit our safety and compliance page at [www.aidoc.com](http://www.aidoc.com).

### References

- 1 Health Ethic & Governance. (2021). Ethics and governance of artificial intelligence for health. <https://www.who.int/publications/i/item/9789240029200>
- 2 Dunn, S. & OWASP. (2024). LLM AI Cybersecurity & Governance Checklist. [https://owasp.org/www-project-top-10-for-large-language-model-applications/llm-top-10-governance-doc/LLM\\_AI\\_Security\\_and\\_Governance\\_Checklist-v1.pdf](https://owasp.org/www-project-top-10-for-large-language-model-applications/llm-top-10-governance-doc/LLM_AI_Security_and_Governance_Checklist-v1.pdf)