

Ethical considerations in the use of Artificial Intelligence (AI) in Healthcare and Washington's approach to Generative AI

Katy Ruckle, JD, FIP State Chief Privacy Officer, WaTech July 24, 2024



Today's Agenda

| ΤΟΡΙΟ | TIME | SPEAKER |
|---------------------------------------|--------------------|---------------|
| Welcome & Introduction | 10:00am – 10:05am | Rhonda Mendel |
| Ethical Issues of Al in Healthcare | 10:05am to 10:25am | Katy Ruckle |
| Al's 'black box' problem | 10:25am to 10:35am | Katy Ruckle |
| WA State's Approach to Al | 10:35am to 10:50am | Katy Ruckle |
| Questions | 10:50am – 11:00am | Katy Ruckle |

Speaker Intro

- Katy Ruckle, JD
- State Chief Privacy Officer
 - Co-chair on Washington's AI Community of Practice.
 - Subcommittee lead on AI Policy Development.
 - Appointed by State Attorney General Bob Ferguson to AGO AI Task Force.





Public Sector Impacts

Media headlines



Another Arrest, and Jail Time, Due to a Bad Facial Recognition Match

SEVEN INDIVIDUALS WITH DISABILITIES CONTINUE THE Care LEGAL FIGHT AGAINST SECRETIVE MEDICAID HOME CARE CUTS





Ethical Issues of AI in Healthcare

Objective:

To discuss and consider the ethical implications of using AI and Generative AI in healthcare settings.

 Importance: As AI and Generative AI technology advances, ethical considerations become increasingly critical to ensure patient well-being and trust.





WaTech Washington Technology Solutions

Definitions

• Artificial intelligence (AI) is machine-based technology that is capable producing an output, including predictions, recommendations or decisions, and uses data and inputs to:

- Assess physical or virtual environments.
- Provide an analysis in a manual or automated manner (such as by using machine learning).
 Use inference to create options for outcomes.
- "Generative AI" is a technology that can create content, including text, images, audio, or video, when prompted by a user. Generative AI systems learn patterns and relationships from large amounts of data, which enables systems to generate new content that may be similar, but not identical, to the underlying training data.



Applications of AI and Generative AI in Healthcare

- * **Medical Imaging:** Use for diagnostics through image reconstruction and interpretation.
- Clinical Decision Support: Assisting healthcare providers in treatment planning and personalized medicine.
- Healthcare Operations: Optimizing workflows and resource allocation such as administrative procedures , making them more efficient, faster, and less expensive.
- Personalized medicine: AI can analyze a patient's medical history, genetic information, and lifestyle to predict disease risks and suggest treatment options.
- * **Digital health:** AI can make medical equipment smarter and help publish results faster. It can also help medical professionals create more efficient diagnosis charts and get insights.
- * **Telemedicine:** Al can help with remote patient monitoring, automated reminders, and generating insights for healthcare providers, especially for chronic diseases.
- * **Public health support:** AI can quickly process large amounts of data and derive insights.



Applications of AI and Generative AI in Healthcare

- Drug discovery and development: AI can accelerate development by analyzing chemical compounds, biological data, and research papers to identify potential new drugs.
- Healthcare Information Retrieval and Natural Language Processing in EHRs: Aids in information retrieval and QI initiatives, addressing the challenge of searching extensive unstructured data.
- Clinical Trial Design Optimization: AI can analyze existing studies, identify relevant patient groups, and identify
 potential outcomes from past results, enabling researchers to optimize trial design, ensure inclusion/exclusion criteria
 are relevant, and reduce costs associated with drug development.
- Patient education and engagement: Tailored communication for patients to actively participate in their healthcare. Generate personalized educational materials, treatment plans, and wellness recommendations based on a patient's specific health status, preferences, and goals.
- Healthcare chatbots and virtual assistants: Conversational agents powered by AI can interact with patients, answer their healthcare-related questions, and provide relevant information on symptoms, conditions, treatments, and preventive measures.
- Medical literature summarization: Al can automatically summarize large volumes of medical literature, research articles, and clinical guidelines into concise summaries, saving time for healthcare professionals.



•

Ethical Challenges

- Privacy and Data Security: Risk to patient data confidentiality and protection against breaches.
- · Bias and Fairness: Biases in training data that could lead to inequitable treatment.
- Automation Bias: Over-reliance on or deference to Al-generated decisions or outcomes.
- Informed Consent: Consent for Al-assisted procedures and diagnoses in a transparent way
 - **Accountability:** Responsibility for Al-generated decisions and outcomes.



Privacy and Data Security

Issue: Al systems require vast amounts of patient data in order to build, train and maintain Al models. Using identifiable patient data to create and maintain generative Al models raises concerns about privacy, unauthorized disclosure or access, and informed consent regarding use of data.

Your Personal Information Is Probably Being Used to Train Generative AI Models

Mitigation:

- · Implementing robust encryption.
- · Anonymization techniques.
- · Strict access controls.



Bias and Fairness

• **Issue:** Al algorithms can perpetuate biases present in training data, leading to disparities in healthcare outcomes.

Media headline

Algorithms Are Making Decisions About Health Care, Which May Only Worsen Medical Racism

Mitigation:

- · Regular auditing of algorithms.
- . Diversifying training datasets.
- . Transparent reporting on performance metrics.



Automation Bias

• **Issue:** Healthcare professionals can be lulled into an over reliance on AI generated medical diagnoses or defer to AI diagnoses instead of relying on their own medical expertise.

• Unique Gen Al Risks:

- Especially compelling quality about it due to its conversant air.
- Sense of confidence Gen AI has in its answers.
- Repeated correctness If clinician finds Gen AI is consistently right, it becomes easier to rely on its predictions.
- Slippery slope that might not quite be the same as other modes of medical knowledge gaining.



2024 Published Case Study re Automation Bias

CASE STUDYf × in ⊡When the Model Trains You: Induced BeliefRevision and Its Implications on ArtificialIntelligence Research and Patient Care — ACase Study on Predicting ObstructiveHydronephrosis in Children



Study Findings re Automation Bias

- "After a silent trial of our hydronephrosis AI model, we observed an unintentional but clinically significant change in practice characterized by a reduction in nuclear scans from 80 to 58%..."
- "This phenomenon occurred in the absence of any identifiable changes in clinical workflow, personnel, practice guidelines, or patient characteristics over time."
- "We postulate that *repeated exposures to model predictors and their corresponding labels led to a change in clinical decision-making* based on a learned intuition of the model's behavior."



Automation Bias

• Mitigation:

- Train and increase awareness about automation bias and its risks
- Encourage a culture of skepticism emphasize critical thinking skills
- Highlight importance of "human" in medical professional-machine collaboration
- Red Teaming Testing results in novel ways
- Diversify perspectives second opinions (from both humans or other Al models)



Informed Consent

• **Issue:** Patients may not fully understand AI's role in their healthcare decisions, potentially compromising autonomy.

SPECIAL REPORT

An invisible hand: Patients aren't being told about the AI systems advising their care



- Educational material and transparency.
- Layman's terms and avoiding jargon.
- Opportunity for questions.
- Patient decision-making process.
- Documentation and consent form.

- Ongoing communication and updates.
- Respecting patient privacy.
- Training and support for healthcare providers.
- Feedback and evaluation.



Educational Material and Transparency:

- Clearly explain why Generative AI is being used, such as improving diagnostic accuracy, personalizing treatment plans, or enhancing medical imaging interpretation.
- Transparently discuss any potential risks or limitations associated with AI use
- Help patients understand how AI compares to traditional methods of diagnosis or treatment planning.

• Layman's Terms and Avoiding Jargon:

- Present information in a language that patients can easily understand, avoiding technical jargon that might confuse or overwhelm them.
- Use analogies or real-world examples to illustrate how AI might assist in their specific healthcare scenario.

Opportunity for Questions:

- Provide patients with ample opportunity to ask questions and seek clarification about AI's role in their care.
- Encourage open dialogue to address concerns or misconceptions patients may have.



Patient Decision-Making Process:

- Ensure patients have enough time to consider the information provided before making a decision.
- Respect patients right to decline AI-assisted procedures or diagnostics if they are uncomfortable or have specific concerns.

• Documentation and Consent Form:

- Develop a consent form outlining the use of Generative AI in healthcare.
- Clearly outline what data will be used, how AI will be employed, and any foreseeable implications.
- Ensure the consent form is written in clear, plain language and covers all pertinent aspects discussed during the informed consent process.

Ongoing Communication and Updates:

- Throughout the treatment or diagnostic process, maintain open communication with patients about any developments or changes related to AI use.
- Provide updates on the accuracy and reliability of AI-generated insights, reinforcing trust and transparency.



Respecting Patient Privacy:

 Emphasize the measures taken to protect patient privacy and confidentiality when using AI, such as data encryption and anonymization techniques.

Training and Support for Healthcare Providers:

- Ensure healthcare providers are adequately trained to explain AI technologies and their implications to patients.
- Provide resources or guidelines to assist providers in conducting effective informed consent discussions.

Feedback and Evaluation:

- Encourage patients to provide feedback on their experience with AI-assisted healthcare.
- Use patient feedback to continuously improve the informed consent process and address any concerns that arise.



Accountability

• **Issue:** Unclear accountability for AI-generated decisions can complicate liability and responsibility in case of errors or harm.

Doctors are liable for the AI they use, state medical board group says



BY: RUTH READER | 05/02/2024 05:45 PM EDT





Accountability



Source: National Library of Medicine article published 11-27-23

- Mitigation:
- Establishing clear guidelines for human oversight.
- Decision-making transparency
- Legal frameworks.
- Ensuring informed consent
- Adopting Al governance in ethical principles



Al's "black box" problem

What do we mean by "black box" problem?



- The "black box" problem in the context of AI refers to the opacity or lack of transparency in understanding how AI systems arrive at their decisions or outputs.
- This issue is particularly pertinent in complex AI models, including deep learning neural networks, where the internal workings are highly intricate and not easily interpretable by humans.



What are we talking about?

| ARTIFICIAL INTELLIGEN Human intellig | NCE | machines | MACH LEAR | HINE NING | ificial intelligence | DEEP LEARNING | | |
|--|--|-------------|-------------------------------------|---|----------------------|--|---|--|
| | | | | | | A technique for implementing machine learning | GENERATIVE AI AI that can produce new content (e.g. text, images, video) | |
| 1950 Technology intelligence predictive a | 1960 y that emulates hum e, perception, and abilities | 1970 nan | 1980 U (e n d a p | 1990 sing algorithms e.g Bayesian etworks, neural etworks) to parse ata, learn from it, nd determine or redict something | 2000 | 2010 Neural networks with "deep" layers that can be trained with massive amounts of data | 2020 Generative AI le patterns and rel from massive ar data, which ena generate new co may be similar, not identical, to underlying train | arns ationsh nounts bles th ontent but the ing da |

27



Al and Gen Al "Black Box" Characteristics

• Complexity of AI Models:

- Many AI systems, especially those based on deep learning techniques, involve numerous layers of interconnected nodes (neurons) that process data through nonlinear transformations.
- Models can have millions of parameters, making it challenging to decipher how each parameter contributes to the final decision.

• Non-Interpretable Features:

- AI models often operate on high-dimensional data and extract abstract features that may not have clear humaninterpretable meanings. For instance, in image recognition, a deep learning model might identify patterns that are important for classification but are not easily understandable in traditional terms.
- Inaccessible Decision-Making Process:
 - Unlike traditional algorithms where decisions are based on explicit rules or logic, AI models derive conclusions based on statistical patterns learned from vast amounts of data.
 - The decision-making process is distributed across layers of the model, making it difficult to pinpoint which factors influence a specific output.



Al and Gen Al "Black Box" Characteristics

Implications for Trust and Accountability:

• The opacity of AI decision-making raises concerns about trust and accountability. For critical applications such as healthcare, finance, or autonomous vehicles, stakeholders (patients, regulators, users) need assurance that decisions are sound and can be interpreted or explained.

Bias and Fairness:

• The lack of transparency can exacerbate issues related to bias in AI systems. If biases are present in the training data or inadvertently introduced during model development, it can be challenging to detect and mitigate them without understanding how decisions are made.

Regulatory and Audit Challenges:

 Regulators and ethical guidelines increasingly require transparency and accountability in AI systems, particularly when they impact individuals' lives or fundamental rights. Without transparency, it becomes difficult to audit AI systems to ensure compliance with regulations or ethical standards.

Guiding Principles for Generative Al Use



Safe, secure, and resilient:

- Al should be used with safety and security in mind, minimizing potential harm and ensuring that systems are reliable, resilient, and controllable by humans.
- Al systems used by state agencies should not endanger human life, health, property, or the environment.

• Valid and reliable:

 Agencies should ensure AI use produces accurate and valid outputs and demonstrates the reliability of system performance.



• Fairness, inclusion, and nondiscrimination:

- Al applications must be developed and utilized to support and uplift communities, particularly those historically marginalized.
- Fairness in AI includes concerns for equality and equity by addressing issues such as harmful bias and discrimination.

Privacy and data protection:

- Al should be used to respect user privacy, ensure data protection, and comply with relevant privacy regulations and standards.
- Privacy values such as anonymity, confidentiality, and control generally should guide choices for AI system design, development, and deployment.
- Privacy-enhancing AI should safeguard human autonomy and identity where appropriate.



• Transparency and auditability:

- Acting transparently and creating a record of AI processes can build trust and foster collective learning.
- Transparency reflects the extent to which information about an AI system and its outputs is available to the individuals interacting with the system.

Accountability and responsibility:

 As public stewards, agencies should use generative Al responsibly and be held accountable for the performance, impact, and consequences of its use in agency work.



• Explainable and interpretable:

- Agencies should ensure AI use in the system can be explained, meaning "how" the decision was made by the system can be understood.
- Interpretability of a system means an agency can answer the "why" for a decision made by the system, and its meaning or context to the user

• Public purpose and social benefit:

• The use of AI should support the state's work in delivering better and more equitable services and outcomes to its residents.



Washington's Approach to Al



Al Initiatives in State of Washington

AI CoP

- Steering Committee
- State Agencies
- Local Gov't
- Subcommittees
 - Risk
 - Policy
 - Use cases
 - Local Gov't

EO 24-01

- Deliverables
 - WaTech
 - DES
 - OFM
 - 00E
 - WTB

AGO AI Taskforce

- Executive Committee
 - (19 members)
 - Industry, advocacy, government, legislative members
- Subcommittees



AI CoP

- Governance Structure
 - Representation from WaTech, State Agency, and Local Government

Steering Committee Objectives

- Develop a set of **guidelines** and **policies**
- Identify and document best practices
- Establish a **governance structure** and develop mechanisms for accountability and oversight
- **Document use cases** and examine potential societal impact
- Facilitate collaboration and knowledge sharing
- Promote alignment of new AI technologies to business and IT strategies





DRAFT Gen AI Policy Sections for Consideration

- Responsible Al
- Al Contact
- Al Inventory
- Risk Assessments
- Training
- Public models

- DSA requirements/Contract Terms
- Accuracy and Transparency
- High Risk GenAl Use
- GenAl Planning
 - Data Quality
 - Monitoring



Executive Order 24-01 on Al

| EO SEC. | LEAD AGENCY | NAMED COLLABORATORS | DELIVERABLE | DEADLINE |
|------------|------------------|---|---|----------------|
| 2 | WaTech | Cabinet Agencies | Report of Gen AI initiatives for agencies | September 2024 |
| 3 | WaTech | DES | Initial Guidelines for Procurement | September 2024 |
| 4 | DES | Office of Equity WaTech | Training plan for state workers | January 2025 |
| 5 | WaTech | Office of Equity Community members Tribal governments SMEs State agencies | Guidelines on impact of adopting Gen AI on vulnerable communities | December 2024 |
| 6 | DES | WaTech | Contract terms templates | January 2025 |
| 7 | Office of Equity | WaTech DES WTB | Accountability Framework | September 2024 |
| 8 | WaTech | | Risk assessments | December 2024 |
| 9 | OFM | Labor organizations WTB WaTech | Report of impact of Gen AI on state workforce | December 2024 |
| 10 | WTB | | Identify and create research opportunities | January 2025 |



AGO Task Force Subcommittees

Education and Workforce Development

Public Safety and Ethics

Healthcare and Accessibility

State Security and Cybersecurity

Government and Public Sector Efficiency

Consumer Protection and Privacy

Labor

Industry and Innovaction



Additional References and Information

- <u>Defining medical liability when artificial intelligence is applied on</u> <u>diagnostic algorithms: a systematic review</u>
- <u>Case Study on Predicting Obstructive Hydronephrosis in Children</u>
- <u>Dissecting racial bias in an algorithm used to manage the health of</u> <u>populations</u>
- Navigating the Responsible and Ethical Incorporation of Al into <u>Clinical Practice</u>



OPDP trainings and resources at **Government Agency Resources** on our webpage.

www.watech.wa.gov/privacy

privacy@watech.wa.gov



Questions?



43