BC 6.935 - Use of Artificial Intelligence (AI) and Generative AI (GenAI) Policy

I. POLICY:

Baystate Health (BH) recognizes the importance of AI and GenAI in the development of new tools and solutions in healthcare. BH also recognizes that AI technologies pose risks that can negatively impact our patients and our organization. Information is shared with a GenAI tool by user prompts, or a series of instructions or questions for the tool. Generally, providing access to information constitutes sharing data with the tool. The sharing of data potentially makes confidential or sensitive information public as the tool may train its model on the data shared. In some cases, data that has been anonymized could be linked to personal information and become exposed. Any patient or employee personal information, company proprietary information or intellectual property, or otherwise confidential information entered into the prompt may appear in other users' output.

Therefore, it is the policy of Baystate Health that any product or solution used or developed at BH that utilizes AI or GenAI will be trustworthy (valid and reliable), safe, secure and resilient, transparent, privacy-enhanced, and fair (with Harmful AI Bias managed).

Systems built on AI or GenAI will be deployed only after an adequate evaluation of its purposes and objectives, its benefits, as well as its risks, and after review by the BH Information Security department (BHIS). Non-Approved, publicly available AI may be used in the limited circumstances described in this policy but may never be used to diagnose a patient, recommend a particular product or therapy or to otherwise substitute for the clinical judgment of a qualified healthcare professional.

BH will not discriminate on the basis of race, color, national origin, sex, age, gender, gender identity or expression, sexual orientation or disability in its health programs or activities through the use of AI or patient care decision support tools.

The procedures in this policy must be followed when using AI or GenAI tools or products.

II. PURPOSE:

- To foster the legal, ethical, and responsible design, development, deployment, and use of AI and GenAI technology and to mitigate the risks associated with the use of AI and GenAI.
- To provide guidance on acceptable and prohibited uses and employee obligations related to the use or development of AI or GenAI models and applications.
- To ensure that BH complies with applicable privacy laws when using AI or GenAI products.

III. SCOPE:

This policy applies to all employees, vendors, contractors, or business associates who have access to BH patient clinical, or business information stored on its computers or have access to its computer resources or network. Devices include workstations, servers and laptops which are used to directly or remotely access the BH network, and those that use AI or GenAI applications or development of GenAI models, including internal models, third party models, or publicly available applications such as ChatGPT. This policy applies to the use of those applications on company devices and/or personal devices when used for work purposes.

IV. DEFINITIONS:

Artificial Intelligence (AI): AI refers to the theory and development of computer systems able to perform tasks normally requiring human intelligence, in order to deliver solutions that can automate routine tasks, draw data-based insights, or augment human activities. (U.S. Department of Health and Human Services Artificial Intelligence (AI) Strategy January 2021) AI-assisted devices include virtual reality, 5G, cloud computing technology, big data analysis, wearable devices, the Internet of Things, etc.

Chatbot: A computer program designed to simulate conversation with a human user, usually over the internet; especially one used to provide information or assistance to the user as part of an automated service.

Generative AI (GenAI) refers to an artificial intelligence technology that synthesizes new versions of text, audio, or visual imagery from large bodies of data in response to user prompts. GenAI models can be used in stand-alone applications or incorporated into other applications such as internet search engines or word processing applications. Examples of Publicly Available AI includes OpenAI, ChatGPT, Bard/Gemini, Google MedLM, Mayo Clinic's Open Evidence.

Hallucinations: Fictitious answers created by GenAI tools. AI hallucinations can take many different forms. Some common examples include:

- Incorrect predictions: An AI model may predict that an event will occur when it is unlikely to happen. For example, an AI model that is used to predict the weather may predict that it will rain tomorrow when there is no rain in the forecast.
- False positives: When working with an AI model, it may identify something as being a threat when it is not. For example, an AI model that is used to detect fraud may flag a transaction as fraudulent when it is not.
- False negatives: An AI model may fail to identify something as being a threat when it is. For example, an AI model that is used to detect cancer may fail to identify a cancerous tumor. (GoogleCloud)

Harmful AI Bias: systematic and repeatable errors in AI systems that create unfair outcomes, such as placing privileged groups at systematic advantage and unprivileged groups at systematic disadvantage. AI systems can reinforce harmful bias when trained on prejudiced or unrepresentative data.

Patient care decision support tool means any automated or non-automated tool, mechanism, method, technology, or combination thereof used by a covered entity to support clinical decision-making in its health programs or activities.

***Predictive Decision Support Intervention (Predictive DSI)*:** technology that supports decision-making based on algorithms or models that derive relationships from training data and then produce an output that results in prediction, classification, recommendation, evaluation, or analysis. 89 FR 1192 (codified at 45 CFR 170.102).

V. REQUIREMENTS FOR THE USE OF AI OR GenAI

All products or systems using AI or GenAI must meet the following requirements:

A. Privacy, Security and Confidentiality:

All AI or GenAI tools or products must be approved by BHIS prior to use, except where otherwise specifically allowed in this policy. BHIS will review the product to ensure that it is safe and secure and that the sharing of any patient or business information does not violate any BH privacy or information security policies. BHIS will also review to ensure that the AI system can maintain confidentiality, integrity, and availability through protection mechanisms that prevent unauthorized access and use (NIST AI 100-1). BHIS will also check the reputation of the tool developer and any third-party services used by the tool.

To maintain the security of our data and IT systems and the confidentiality of our company's sensitive information, employees may not:

- Attempt to gain access to unapproved GenAI applications when using BH systems or networks, conducting business on behalf of BH, or when using BH data.
- Use Baystate credentials, email addresses, or telephone numbers as a login to publicly available GenAI applications, such as ChatGPT.
- Install non-approved Application Programming Interfaces (APIs), plug-ins, connectors, or software related to GenAI systems.
- Implement or use in any way unsupervised or unvalidated code generated by GenAI on BH systems.
- Use GenAI tools on personal devices to circumvent company policies or safeguards.
- Enter BH intellectual property, personal information (PII) of employees, patients (PHI), or other third parties (such as names, dates of birth, SSNs) into any non-approved or publicly available GenAI application, such as ChatGPT.

B. Transparency

The use of AI should be transparent. To maintain transparency with patients and others, employees must:

- Clearly attribute any output used for work purposes to the GenAI application that created it through a footnote or other means visible to the reader.
- Maintain a record of GenAI used for work purposes and be able to share those records with your manager or other authorized company personnel upon request.
- When using GenAI to interact with a person, clearly and conspicuously disclose to the person with whom the GenAI interacts, that the person is interacting with GenAI and not a human.
- Obtain patient consent when using AI to record a patient interaction.

C. Verification of AI Output:

Many open-source GenAI models are trained on large, publicly available datasets (e.g., through data extraction of public webpages). The outputs may therefore contain copyrighted information, false or fraudulent information, or others' intellectual property. GenAI tools may be useful but are not a substitute for human judgement and creativity. Anyone using an GenAI platform that has been approved by BHIS for clinical or business purposes must:

- Verify the output prior to use.
- Validate the accuracy of the data.

- Have an appropriate BH clinician review, correct, and approve the output of any clinical documentation that is created using approved AI or GenAI prior to adding it to a patient's medical record.
- Not use any output that contains material subject to copyright protection in any internal or externally facing materials.
- Not use non-approved AI to provide certainty with respect to a diagnosis, to recommend a particular product or therapy or to otherwise substitute for the clinical judgment of a qualified healthcare professional.
- Not use non-approved AI to diagnose or treat users.

D. Managing Harmful AI Risks (including Bias and Hallucinations):

Outputs created by GenAI tools may provide fictitious answers, sometimes referred to as hallucinations. AI tools also incorporate any biases of the data sets that were used to train them. This modeling bias may not always align with Baystate's core values and our commitment to diversity, equity and inclusion. Model output may make systematic errors or favor certain groups, leading to unfair or discriminatory outcomes. Using output from GenAI tools without reviewing it for accuracy or Harmful AI Bias places the company at risk and may harm BH's reputation with patients, the community, and employees. Therefore, to protect our employees and patients from harm, and to protect BH from reputational damage, employees must:

- Make reasonable efforts to identify AI tools that employ input variables or factors that measure race, ethnicity, color, national origin, ancestry, age, gender, gender identity or expression, sex, sexual orientation, or disability and to mitigate the risk of discrimination resulting from the tool's use. This includes algorithmic decision-making that relies on or uses discriminatory inputs and that produces discriminatory results, such as those that have the purpose or effect of disfavoring or disadvantaging a person or group of people based on a legally protected characteristic.
- Review output for Harmful AI Bias and to make sure it meets the standards in BH's Code of Conduct and non-discrimination policies. GenAI-created content that is inappropriate, discriminatory or otherwise harmful to our employees, patients or clinicians must not be used for work purposes. Such use will result in disciplinary action, up to and including termination.
- Review output for Hallucinations.
- Verify that vendors or third parties using AI in their products have controls and processes in place to protect against Harmful AI Bias and Hallucinations.
- Verify the data to ensure no accidental misuse or false or fraudulent statements.
- Not use GenAI applications to create text, audio, or visual content for purposes of committing fraud or to misrepresent an individual's identity.

E. Patient Consent:

If using approved AI technology to record a patient interaction, consent must be obtained from any person whose voice will be recorded. In addition, patient data sent by BH to outside organizations looking to use that data to develop new AI tools, train new GenAI models, or improve the user experience may require consent from the individual whose data will be used. Contact the BH Compliance and Privacy Office for questions or assistance with consent development.

F. Safety/Use of Reputable AI Tools:

Use of AI systems should not "lead to a state in which human life, health, property, or the environment is endangered" (Source: ISO/IEC TS 5723:2022).

Employees should only use reputable AI tools and should be cautious when using tools developed by individuals or companies without established reputations.

VI. MONITORING:

Managers are responsible for:

- Assisting staff with escalating questions on use of GenAI to BHIS or Compliance.
- Assuring that GenAI-generated content is appropriately labeled as such in all media.

BH reserves the right to access and monitor the use of GenAI applications on any company-issued devices or that appear on company managed networks to ensure compliant use of these systems. Users who fail to comply with any provision of this policy may be subject to discipline up to and including termination of employment. Violations by contractors may be considered breach of contract and result in removal from assignment. If monitoring systems and processes detect a possible policy violation or if a User reports a possible policy violation, the suspect event should be processed using appropriate security incident response processes (see the Cybersecurity Incident Response Policy).

VII. REFERENCES:

89 FR 37522 Nondiscrimination in Health Programs and Activities

National Institute of Standards and Technology (NIST) Artificial Intelligence Risk Management Framework (AIRMF1.0) https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf NIST AI 100-1 (2023), *https://doi.org/10.6028/NIST.AI.100-1*

October 30, 2023 Biden-Harris Executive Order E.O. 14110 "Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence" https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/

Department of Health and Human Services OMB M-21-06 (Guidance for Regulation of Artificial Intelligence Applications) https://www.hhs.gov/sites/default/files/department-of-health-and-human-services-omb-m-21-06.pdf

"Health Data, Technology, and Interoperability: Certification Program Updates, Algorithm Transparency, and Information Sharing" (HTI-1) Final Rule 45 C.F.R. Parts 170 and 171

U.S. Dep't of Health & Hum. Servs., Off. of the Nat'l Coordinator for Health Info. Tech., *Health Data, Technology, and Interoperability: Certification Program Updates, Algorithm Transparency, and Information Sharing,* Final Rule, 89 FR 1192 (January 9, 2024);

Consumer Fin. Protection Bureau, U.S. Dep't of Justice, U.S. Equal Employment Opportunity Comm'n, & Fed. Trade Comm'n, *Joint Statement on Enforcement Efforts Against Discrimination and Bias in Automated Systems* (Apr. 2023), *https://files.consumerfinance.gov/f/documents/cfpb_joint-statement-enforcement-against-discrimination-bias-automated-systems_2023-04.pdf;*Am. Med.

Ass'n, *American Medical Association Principles for Augmented Intelligence Development, Deployment, and Use,* pp. 2-4 (2023), *https://www.ama-assn.org/system/files/ama-ai-principles.pdf*.

MA Attorney General Advisory on the Application of the Commonwealth's Consumer Protection, Civil Rights, and Data Privacy Laws to Artificial Intelligence.

VIII. CROSS REFERENCES:

HR 802 - Confidentiality
HR 804 - Corrective Action

APPROVALS

		Date:
Approved:	BH Policy Management Committee Data Governance Executive Committee	7/16/2024 7/17/2024
Policy Owner: (department name)	Information Security	
Revising Author:	Tina Basch, VP and Chief Information Security Officer	
Reviewers: (Stakeholders)	Renée Wroth, VP and Chief Compliance & Privacy Officer Data Use Protection Committee	
Replaces:		
Original Effective Date:		7/16/2024
Review Dates:		
Revision Dates:		

Appendix A

Examples of Acceptable Uses	***Examples of Not Acceptable Uses***
Using ChatGPT to generate generic business content such as website content or marketing materials	Using Publicly Available AI such as ChatGPT, Bard/Gemini, Claude 2, MedLM to create clinical medical record documentation
Using Baystate approved AI tools, such as Nuance, DAX	Entering patient information into a non-approved Publicly Available AI platform such as ChatGPT
Using MedLM for educational purposes such as engaging in non-patient specific medical questioning and answering	Using MedLM for any diagnostic or therapeutic purpose, or in direct patient care.
Using Publicly Available AI for clinical reference purposes	