San Francisco Department of Public Health

Policy & Procedure

Policy & Procedure Title: Artificial Intelligence (AI) Policy	
Category: Information Systems	
Effective Date: August 1, 2024	**Last Reissue/Revision Date:** n/a
DPH Unit of Origin: SFDPH Information Governance Steering Committee	
Policy Contact - Employee Name and Title; and/or DPH Division: Eric Raffin, Chief Information Officer	DocuSigned by: *Eric Raffin* 9BECFDD2B627406…
Contact Phone Number(s): (628) 271-6020	
Distribution: DPH-wide ☒	**If not DPH-wide, other distribution:**

1.   Purpose

This policy provides guidelines and establishes standards and procedures for the use of artificial intelligence (AI) at the San Francisco Department of Public Health (SFDPH). This policy applies to anyone working at SFDPH, including DPH staff, volunteers, and contractors. AI systems are tools, neither intrinsically good nor bad. We know there is a history of bias and inequity in how the data that AI tools draw from are created. Healthcare algorithms and AI bias can contribute to existing health disparities for certain populations based on race, ethnicity, gender, age, or other demographic factors. To avoid structural bias in AI systems, this policy addresses the need for intentional and holistic review of AI systems and the data they draw from.

2.   Policy

It is the policy of SFDPH to ensure that the use of AI tools complies with current and emerging federal, state, and local laws and regulations and conforms to the ethical principles for use of technologies in healthcare and public health settings.

In recognition of the transformative potential of AI in healthcare, this policy establishes comprehensive guidelines and principles for the ethical and responsible development, deployment, and governance of AI technologies within SFDPH. It draws upon the principles outlined in the United Nations Declaration of Human Rights (UDHR), specifically Article 3 emphasizing the right to both physical and mental security, alongside the World Health Organization's Ethics and Governance of Artificial Intelligence for Health. By adhering to the principles outlined in this policy, SFDPH aims to harness the potential of AI to improve health outcomes, enhance the service experience, and promote equity and justice in healthcare while upholding the fundamental rights and dignity of individuals. This policy provides a comprehensive framework for the ethical and responsible development, deployment, and governance of AI technologies at SFDPH, ensuring that they serve the best interests of patients, communities, and SFDPH.

The mission of the San Francisco Department of Public Health is to protect and promote the health of all San Franciscans.
We shall ~ Assess and research the health of the community ~ Develop and enforce health policy ~ Prevent disease and injury ~
~ Educate the public and train health care providers ~ Provide quality, comprehensive, culturally proficient health services ~ Ensure equal access to all ~

3. Definitions

A. Artificial intelligence (AI): a group of technologies that can perform complex cognitive tasks that typically require human intelligence. This includes, but is not limited to:

   a. Generative AI (Gen-AI): Subset of AI that focuses on creating models capable of generating new, original content. Gen-AI is trained on vast amounts of data from the internet and other sources. Examples include products that create AI-generated text or images based on user prompts.

   b. Enterprise AI: Subset of AI used by organizations to improve operations, leveraging data generated within the organization. Examples include chatbots, cognitive computing, predictive analytics, process automation, and personalization.

   c. Narrow AI: Subset of AI that operates under a limited set of constraints and is task-specific, following pre-programmed rules or learned patterns from data. Examples include virtual assistants, search engines, and recommendation engines.

   d. Language Models: Specialized type of AI trained using vast amounts of data to enable it to recognize, translate, predict, or generate content. Language models are used for text generation, language translation, crafting creative content, answering queries, and code generation.

   e. Machine Learning: Type of AI algorithm that lets software applications become more accurate at predicting outcomes autonomously. Examples include facial recognition, credit card fraud detection, and spam detection.

B. Patient: Any individual who is a current or previous SFDPH patient, client, or resident who has received, is receiving, or will receive healthcare or public health services from SFDPH.

C. Protected Health Information (PHI): Individually identifiable health information maintained or transmitted in any medium: oral, written, or electronic. It involves any information that identifies an individual AND relates to: 1. The individual's past, present or future physical or mental health; OR 2. The provision of health care to the individual; Or 3. The past, present or future payment for health care.

4. Principles

A. *Human Rights and Dignity:* All AI initiatives within the SFDPH must respect and uphold the principles of human rights and dignity as enshrined in Article 3 of the United Nations Declaration of Human Rights. This includes the right to life, liberty, and security of person, encompassing both physical and mental well-being. Healthcare AI systems must prioritize the preservation of individual autonomy, privacy, and dignity in all interactions and interventions.

B. *Beneficence and Non-Maleficence:* AI technologies deployed in healthcare must prioritize the well-being of patients, clients, staff, and communities. They should aim to improve health outcomes, minimize harm, and support equitable access to healthcare services for all individuals, regardless of socioeconomic status or background. AI algorithms and interventions should be designed to optimize patient/client safety, minimize errors, and mitigate potential risks associated with their use. The primary objective of beneficial AI systems must be to maximize the realization of human preferences, which means the behavior of an AI system must align with human behavior.

C. *Transparency and Accountability:* Leaders, policymakers, developers, and healthcare providers must ensure transparency in the design, development, and deployment of AI systems. They should be accountable for the decisions made by AI algorithms, and mechanisms for recourse and redress should be established in case of errors, biases, or adverse outcomes. Responsibility can be assured by application of human warranty, which implies evaluation of AI solutions by staff and patients/clients in the development, deployment, and use of AI technologies. AI models and systems should be intelligible and explainable to all parties involved with the use of an AI system, meaning that the models in use are explainable in common language such that a person can understand how an AI system reached a conclusion, decision, or action, and from what data. Clear documentation of AI algorithms, data sources,

and decision-making processes will be provided to facilitate external scrutiny and accountability. Users of AI systems shall clearly indicate when and how AI has been employed to create or enhance content. This transparency helps maintain the integrity of the work, allowing audiences to differentiate between human-created and AI-generated material.

D. ***Equity and Justice:*** AI initiatives should strive to reduce health disparities and address systemic inequalities in healthcare delivery in concert with SFDPH strategic objectives. They should prioritize the needs of marginalized, underserved, under-represented, and under-recognized populations, ensuring that AI technologies do not exacerbate existing disparities but rather contribute to closing gaps in healthcare access and outcomes. Efforts should be made to ensure fair distribution of AI-enabled healthcare resources and services, with strong consideration for social and environmental determinants of health. When evaluating AI systems, efforts should also be made to verify that a holistic evaluation tool has been used to review the fairness and consistency of language models. Decision-making about the use of AI should include diverse representation, including those representing patient and community viewpoints whenever possible.

E. ***Autonomy and Informed Consent:*** Patients have the right to make informed decisions about their healthcare. AI applications should support patient autonomy by providing clear and understandable information about how their data will be used, enabling them to give informed consent for its collection, storage, and analysis. Patients should have the opportunity to opt-out of AI-driven interventions and have control over the sharing of their health information in accordance with federal, state, and local regulations. Meaningful engagement with patients/clients, communities, advocacy groups, and other stakeholders shall be prioritized throughout the AI lifecycle, from conception to implementation, to ensure their perspectives and concerns are considered. Participatory approaches such as co-design and community consultation should be employed to foster trust, transparency, and collaboration between AI developers, healthcare providers, and the public.

F. ***Data Privacy and Security:*** AI systems must adhere to the highest standards of data privacy and security. Patient/client data should be collected, stored, and processed in compliance with relevant regulations and ethical guidelines. Measures should be implemented to safeguard against unauthorized access, misuse, or breaches of sensitive and protected health information. Encryption, anonymization, and data minimization techniques should be employed to protect patient/client privacy.

G. ***Continuous Evaluation and Improvement:*** The performance and impact of AI technologies in healthcare should be continuously monitored, evaluated, and improved using defined measurable objectives. This includes assessing their efficacy, safety, and ethical implications, as well as soliciting feedback from stakeholders to inform iterative refinement and optimization, aligning with SFDPH's Lean process improvement program. New AI interventions should be accompanied by training and education on relevant ethical considerations, regulatory requirements, and best practices, as designed by the operational owner sponsoring the AI initiative. Regular audits and assessments should be conducted to identify and address potential biases, errors, or unintended consequences of AI systems.

H. ***Regulatory Compliance:*** The regulatory landscape evolves quickly in the context of emerging technologies. SFDPH needs to ensure that SFDPH use of AI conforms with evolving regulations, including those related to copyright law and intellectual property law governing plagiarism.

5. Procedures

The procedure for the acquisition, use, and evaluation of new AI initiatives is as follows.

A. All AI tools must be acquired in accordance with the SFDPH's standard IT Governance process which requires, at a minimum, privacy, information security, and digital accessibility reviews.

B. AI tools that are already available within existing SFDPH systems will be reviewed to ensure alignment with the principles of beneficial AI outlined in this policy.

C. Requests for operational and IT resources to obtain, implement, or analyze potential new AI tools will follow existing request procedures in alignment with operational and budgetary prioritization. This includes (but is not limited to) the standard IT procurement, project management, Epic enhancement, or information governance request procedures.

D. Additional questions to assess tools for principles listed in this policy will be added to that evaluation process for AI-related requests, including (but not limited to) questions about:

      a. Stakeholder engagement
      b. Disclosure and informed consent
      c. Equity impact assessment
      d. Security and privacy protections
      e. Continuous improvement plan for the larger project incorporating AI-related request
      f. Training and communication plans
      g. Compliance with regulatory requirements related to AI

E. Metrics for AI-related intervention evaluation should be:
      a. Specific (defined in relation to operational or clinical objectives with a defined implementation plan)
      b. Measurable (measurement of benefits and potential consequences)
      c. Aligned (with defined strategic objectives, inclusive and with buy-in of the involved stakeholders)
      d. Realistic (changes that the proposed solution will work as promised and have impact on the intended processes)
      e. Transformative (will use of AI have incremental or transformative impact)

F. Metrics for AI clinical model evaluation should include evaluation of:
      a. Safety/risk, including safety categorization, on-/off-label use under Food and Drug Administration (FDA) regulation, identification and mitigation of potential harms, goal to maintain or improve standard of care, safety to implement
      b. Accuracy, including comparison with previously studied populations, metrics for evaluating model accuracy, performance compared with other models
      c. Fairness/Bias, including performance with diverse populations and mitigation for potential bias
      d. Evidence, including evaluation of peer-reviewed studies, FDA clearance, available evidence from similar institutions

G. As described in the SFDPH Research Policy, institutional review board review must be completed for all AI-related research projects and should address the ethical considerations in this policy.

H. AI-related requests will undergo review by the SFDPH Information Governance Steering Committee, its designated subcommittee, or its subcommittees' designated representative to assess the potential risks, benefits, and compliance with regulatory and ethical standards. Reviews will consider the principles outlined in this policy, as well as relevant ethical frameworks and guidelines developed by professional associations and regulatory bodies.

6. Compliance

Concerns about regulatory or ethical considerations as outlined in this policy can be reported through existing compliance monitoring options including:

- Patient safety / unusual occurrence reporting for safety considerations

- SFDPH Privacy and Compliance Office for potential privacy or compliance violations

- SFDPH Office of Cybersecurity for information security threats / vulnerabilities

7. References

a. CCSF San Francisco Generative AI Guidelines, November 9, 2023.
b. SFDPH "DPH Privacy Policy"
c. SFDPH "HIPAA Compliance – Privacy and the Conduct of Research" policy
d. Executive Order 1110 "Safe, Secure, and Trustworthy Development and Use of Artificial AI," October 2023.
e. California Privacy Protection Agency "Draft Automated Decision Making Technology Regulations," December 2023.
f. World Health Organization "Ethics and Governance of Artificial Intelligence for Health," 2021.
g. Russel, S. (2019) *Human Compatible*, Viking Press

h. International Medical Device Regulators Forum (IMDRF), "Software as a Medical Device": Possible Framework for Risk Categorization and Corresponding Considerations," 2014.
i. Text generated by ChatGPT, OpenAI, December 18, 2023, https://chat.openai.com/chat
j. Center for Research on Evaluation Models, www.CRFM.edu, Stanford University, 2021