State of North Carolina

North Carolina State Government Responsible Use of Artificial Intelligence Framework

August 2024



Table of Contents

Framework	1
Introduction	1
Purpose	1
Policy	1
Scope and Authority	2
Principles and Practices	2
Requirements	4
AI Inventory	4
AI Risk Assessment	4
Appendix A – Glossary	5
Appendix A – Glossary Appendix B – NIST AI Risk Management Framework	5 6
Appendix A – Glossary Appendix B – NIST AI Risk Management Framework NIST AI RMF Core	5 6 6
Appendix A – Glossary Appendix B – NIST AI Risk Management Framework NIST AI RMF Core Govern	5 6 6 7
Appendix A – Glossary Appendix B – NIST AI Risk Management Framework NIST AI RMF Core Govern Map	5 6 7 9
Appendix A – Glossary Appendix B – NIST AI Risk Management Framework NIST AI RMF Core Govern Map Map	5 6 7 9
Appendix A – Glossary Appendix B – NIST AI Risk Management Framework NIST AI RMF Core Govern Map Map Manage	5 6 7 9 1

Document Information

Department Name	N.C. Department of Information Technology
Owner	Secretary and State Chief Information Officer
Title	North Carolina State Government Responsible Use of Artificial Intelligence Framework
Publication Date	Aug. 21, 2024
Version	1.0

Framework

Introduction

Artificial intelligence (AI) is a broad term used to describe an engineered system where machines learn from experience, adjusting to new inputs, and potentially performing tasks previously done by humans. More specifically, it is a field of computer science dedicated to simulating intelligent behavior in computers. It may include automated decision-making (International Association of Privacy Professionals, Glossary, <u>https://iapp.org/resources/glossary</u>, 2024).¹

The state has leveraged certain AI technologies when building out its analytic capabilities to support improved insights. These technologies have the potential to transform society, drive economic growth, support scientific advancement, and help government serve people more effectively and efficiently. They also pose risks that can negatively impact people, organizations, and society.

The State Chief Information Officer supports the use of AI, where appropriate, to improve government innovation, operations, and services in a manner that benefits the people, fosters public trust, builds confidence in AI, protects our state's values, and remains consistent with all applicable laws.

Opportunities for designing, developing, acquiring, and using AI should be sought to improve state government while carefully considering potential risks and how they could best be assessed and managed.

Purpose

The North Carolina State Government Responsible Use of Artificial Intelligence Framework (Al Framework) is designed to encourage responsible exploration and use of Al to benefit the people of North Carolina, foster public trust and confidence in the use of Al, protect our state's values, and ensure that the use of Al remains consistent with all applicable laws, including those related to privacy, civil rights, and civil liberties.

The AI Framework consists of principles, practices, and guidance to agencies who are trying to reap the benefits of AI while reducing privacy and data protection risks when using specific types of artificial intelligence (AI) and supporting the privacy and protection of sensitive data provided to the state by North Carolinians.

Policy

State agencies must follow the common set of principles outlined in the AI Framework when considering the design, development, acquisition, and use of AI in government. The AI Framework is based on principles for AI that build on the <u>Fair Information Practice Principles</u> adopted by the state in May 2022, as well as privacy and security best practices for the use of AI.²

¹ IAPP's definition provides a high-level summary of AI definitions found in NIST's *The Language of Trustworthy AI: An In-Depth Glossary of Terms (*March 22, 2023).

² AI Framework principles are informed by the White House, Office of Science Technology and Policy, Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People, n.d., <u>https://www.whitehouse.gov/ostp/ai-bill-of-rights/</u>.

Scope and Authority

This framework applies to the use of all AI by State Agencies. State Agencies shall have the same meaning as provided in N.C.G.S. § 143B-1320(a)(17).

The AI Framework applies to all systems that use, or have the potential to use, AI and have the potential to impact North Carolinians' exercise of rights, opportunities, or access to critical resources or services administered by or accessed through the state. This includes all AI designed, developed, acquired, or used by state agencies, unless specifically excluded by applicable law.

The AI Framework applies to both existing and new uses of AI; both stand-alone AI and AI embedded within other systems or applications; AI developed both by the agency or by third parties on behalf of agencies for the fulfilment of specific agency missions, including relevant data inputs used to train AI and outputs used in support of decision making; and agencies' procurement of AI systems or applications.

The AI Framework does not apply to basic AI embedded within common commercial products, such as predictive text in word processors or dynamic route adjustment based on real-time traffic conditions in map navigation systems, while noting that government use of such products must nevertheless comply with applicable law and policy to assure the protection of security, privacy, rights, and state values.

Pursuant to N.C.G.S. § 143B-1376 - Statewide Security and Privacy Standards, the State Chief Information Officer (CIO) is responsible for the security and privacy of all state information technology systems and associated data. The State CIO manages all executive branch information technology security and shall establish a statewide standard for information technology security and privacy to maximize the functionality, security and interoperability of the state's distributed information technology assets, including, but not limited to, data classification and management, communications and encryption technologies.

Nothing in this framework shall be construed to impair or otherwise affect: (i) the authority granted by law to a department or agency, or the head thereof; or (ii) the functions of an agency relating to budgetary, administrative, or legislative proposals. This framework should be implemented consistent with applicable law and subject to the availability of appropriations. It is not intended to, and does not, create any right or benefit, substantive or procedural, enforceable at law or in equity by any party against the State of North Carolina, its agencies, or entities, its officers, employees, or agents, or any other person.

Principles and Practices

State government should use AI to support operations to benefit North Carolinians and the public good. Agencies should consider AI in instances where it can help further the agency's mission, enhance service delivery, and improve efficiency and effectiveness. The overarching goal for state government in exploring and using technology, including technology that includes AI, should always be to benefit the people of North Carolina.

These seven principles and associated practices form a blueprint of ethical behavior to guide the state in using AI responsibly to harness its benefits to serve the public while minimizing potential harm. Agencies need to ensure that their AI applications are regularly tested against these principles.

Mechanisms should be maintained to modify, supersede, disengage, or deactivate existing applications of AI that demonstrate performance or outcomes that are inconsistent with their intended use or these principles.

The principles and associated practices are:

- Human-centered: Human oversight is required for all development, deployment, and use of AI. The state should use AI to benefit North Carolinians and the public good. Human oversight should ensure that the use of AI does not negatively impact North Carolinians' exercise of rights, opportunities, or access to critical resources or services administered by or accessed through the state.
- 2. **Transparency and Explainability:** When AI is used by the state, the user agency shall provide notice to those who may be impacted by its use. This notice should identify the use of an automated system, explain why it is used, and how this use contributes to outcomes that impact individuals. This notice should be accessible and written in plain language. Notice should include clear descriptions of the data, the role automation plays in decision-making, and the ability to trace the cause of possible errors.
- 3. **Security and Resiliency:** Systems utilizing AI must undergo pre-deployment testing, risk identification and mitigation, and ongoing monitoring that demonstrates the systems are safe and effective, in keeping with standards for security review for all technology implemented within state government. Systems need to be assessed for resilience to attack, adherence to security standards, and alignment with general safety, accuracy, reliability, and reproducibility.
- 4. **Data Privacy and Governance:** Any use of AI by the state must maintain the state's respect for individuals' privacy and its adoption of the <u>Fair Information Practice Principles</u> throughout the AI lifecycle (development, testing, deployment, decommissioning). This means that privacy is embedded into the design and architecture of IT and business practices. Preservation of privacy should be the default and access to data should be appropriately controlled. Individuals developing or deploying AI systems should be conscious of the quality and integrity of data used by those systems.
- 5. **Diversity, Non-discrimination, and Fairness:** Al should be developed with consultation from diverse communities, stakeholders, and domain experts to identify concerns, risks, biases, and potential impacts of the system. Al needs to be developed to be equitable and control for biases that could lead to discriminatory results. Al systems should be user centric and accessible to all people.
- 6. **Auditing and Accountability:** Users of AI must be accountable for implementing and enforcing appropriate safeguards for the proper use and functioning of their applications of AI, and shall monitor, audit, and document compliance with those safeguards. Agencies shall provide appropriate training to all agency personnel responsible for the design, development, acquisition, and use of AI.
- 7. **Workforce Empowerment:** Staff are empowered in their roles through training, guidance, collaborations, and opportunities that promote innovation that aligns with state or agency missions and goals. This can help state government make best use of Al tools to reduce administrative burdens on staff where feasible and improve overall public service.

Requirements

To properly identify and assess opportunities and risks related to AI, the state must have a comprehensive inventory of AI tools and follow a common framework for risk assessment.

Al Inventory

Agencies must keep an inventory of the tools or applications using AI (including the types of AI) being used, by whom, and for what purposes.³

Agencies' inventories should be reported to the North Carolina Department of Information Technology for transparency and updated, at a minimum, through the annual application portfolio management process. Decisions about AI use should include considerations of continuity.

AI Risk Assessment

Agencies must use the NIST AI Risk Management Framework (AI RMF) to assess and manage risk to individuals, organizations, and society associated with AI before deployment and on a continuing basis once AI is deployed.

The NIST AI RMF was developed in collaboration with the private and public sectors and is an essential tool in improving the ability to incorporate trustworthiness considerations into the design, development, use, and evaluation of AI products, services, and systems.

The Enterprise Security and Risk Management Office (ESRMO) and the Office of Privacy and Data Protection (OPDP) will provide guidance concerning risk assessments for AI in enterprise-level platforms, services, and applications. All risk needs to be assessed and documented.⁴

³ The use of an AI inventory aligns with Executive Order 13960 Promoting the Use of Trustworthy AI in the Federal Government establishes principles for the use of AI in the Federal Government, which establishes a common policy for implementing the principles and directs agencies to catalogue their AI use cases. ⁴ A Privacy Threshold Analysis (PTA) is the initial tool used by the state to identify and document risk.

Appendix A – Glossary

Algorithm: A set of computational rules to be followed to solve a mathematical problem. More recently, the term has been adopted to refer to a process to be followed, often by a computer (NIST Glossary of Al Terms, March 2023).

Artificial Intelligence (AI): Artificial intelligence (AI) is a broad term used to describe an engineered system where machines learn from experience, adjusting to new inputs, and potentially performing tasks previously done by humans. More specifically, it is a field of computer science dedicated to simulating intelligent behavior in computers. It may include automated decision-making (International Association of Privacy Professionals, Glossary, <u>https://iapp.org/resources/glossary</u>, 2024). This definition provides a high-level summary of AI definitions found in NIST's *The Language of Trustworthy AI: An In-Depth Glossary of Terms* (March 22, 2023).

Al Prompt: An instruction or command given to an artificial intelligence tool to carry out a task or function.

Confidential Information: Refers to all information about the organization, its operations, clients, or employees that is subject to reasonable efforts by the organization to maintain its confidentiality and that is not typically disclosed by custom or law to people who are not affiliated with the organization but does not qualify as a trade secret.

Generative AI: A kind of artificial intelligence capable of generating new content such as code, images, music, text (ex: ChatGPT), simulations, 3D objects, videos, and so on. It is considered an important part of AI research and development, as it has the potential to revolutionize many industries, including entertainment, art, and design. (*NIST Glossary of AI Terms*, March 2023)

Hallucination: Generated content that is nonsensical or unfaithful to the provided source content (NIST Glossary of AI Terms, March 2023).

Machine Learning: The study or the application of computer algorithms that improve automatically through experience. Machine learning algorithms build a model based on training data in order to perform a specific task, like aiding in prediction or decision-making processes, without necessarily being explicitly programmed to do so (*NIST Glossary of AI Terms*, March 2023).

Personally Identifiable Information (PII): Information that can be used to distinguish or trace an individual's identity – such as name, Social Security number, biometric data records – either alone or when combined with other personal or identifying information that is linked or linkable to a specific individual (e.g., date and place of birth, mother's maiden name, etc.).

Trustworthy AI: Characteristics of trustworthy AI systems include valid and reliable, safe, secure and resilient, accountable and transparent, explainable and interpretable, privacy-enhanced, and fair with harmful bias managed (*NIST Glossary of AI Terms*, March 2023).

Appendix B – NIST AI Risk Management Framework

Risk management is central to responsible development and use of AI technologies.

Responsible AI practices can help align the decisions about AI system design, development, and uses with intended aim and values.... AI risk management can drive responsible uses and practices by prompting organizations and their internal teams who design, develop, and deploy AI to think more critically about context and potential or unexpected negative and positive impacts. Understanding and managing the risks of AI systems will help to enhance trustworthiness, and in turn, cultivate public trust.

(NIST AI RMF 1.0, January 2023, Executive Summary, p.1)

The NIST AI RMF Core is reproduced below to serve as a starting point for AI risk assessment. NIST will continue to update this framework and provide additional guidance. Always consult NIST directly to ensure access to the most current version of the NIST AI RMF to guide risk assessment.

NIST AI RMF Core

The NIST AI RMF Core provides outcomes and actions that enable dialogue, understanding, and activities to manage AI risks. This is the foundation for Working Groups to create the processes to evaluate and mature AI use cases. The graphic below provides a visual understanding of what is involved in AI risk management.



Govern

Govern 1

Policies, processes, procedures and practices across the organization related to the mapping, measuring and managing of AI risks are in place, transparent, and implemented effectively.

Govern 1.1

Legal and regulatory requirements involving AI are understood, managed, and documented.

Govern 1.2

The characteristics of trustworthy AI are integrated into organizational policies, processes, and procedures.

Govern 1.3

Processes and procedures are in place to determine the needed level of risk management activities based on the organization's risk tolerance.

Govern 1.4

The risk management process and its outcomes are established through transparent policies, procedures, and other controls based on organizational risk priorities.

Govern 1.5

Ongoing monitoring and periodic review of the risk management process and its outcomes are planned, organizational roles and responsibilities are clearly defined, including determining the frequency of periodic review.

Govern 1.6

Mechanisms are in place to inventory AI systems and are resourced according to organizational risk priorities.

Govern 1.7

Processes and procedures are in place for decommissioning and phasing out of AI systems safely and in a manner that does not increase risks or decrease the organization's trustworthiness.

Govern 2

Accountability structures are in place so that the appropriate teams and individuals are empowered, responsible, and trained for mapping, measuring, and managing AI risks.

Govern 2.1

Roles and responsibilities and lines of communication related to mapping, measuring, and managing AI risks are documented and are clear to individuals and teams throughout the organization.

Govern 2.2

The organization's personnel and partners receive AI risk management training to enable them to perform their duties and responsibilities consistent with related policies, procedures, and agreements.

Govern 2.3

Executive leadership of the organization takes responsibility for decisions about risks associated with AI system development and deployment.

Govern 3

Workforce diversity, equity, inclusion, and accessibility processes are prioritized in the mapping, measuring, and managing of AI risks throughout the lifecycle.

Govern 3.1

Decision-makings related to mapping, measuring, and managing AI risks throughout the lifecycle is informed by a diverse team (e.g., diversity of demographics, disciplines, experience, expertise, and backgrounds).

Govern 3.2

Policies and procedures are in place to define and differentiate roles and responsibilities for human-Al configurations and oversight of Al systems.

Govern 4

Organizational teams are committed to a culture that considers and communicates AI risk.

Govern 4.1

Organizational policies and practices are in place to foster a critical thinking and safety-first mindset in the design, development, deployment, and uses of AI systems to minimize negative impacts.

Govern 4.2

Organizational teams document the risks and potential impacts of the AI technology they design, develop, deploy, evaluate and use, and communicate about the impacts more broadly.

Govern 4.3

Organizational practices are in place to enable AI testing, identification of incidents, and information sharing.

Govern 5

Processes are in place for robust engagement with relevant AI actors.

Govern 5.1

Organizational policies and practices are in place to collect, consider, prioritize, and integrate feedback from those external to the team that developed or deployed the AI system regarding the potential individual and societal impacts related to AI risks.

Govern 5.2

Mechanisms are established to enable AI actors to regularly incorporate adjudicated feedback from relevant AI actors into system design and implementation.

Govern 6

Policies and procedures are in place to address AI risks and benefits arising from third-party software and data and other supply chain issues.

Govern 6.1

Policies and procedures are in place that address AI risks associated with third-party entities, including risks of infringement of a third party's intellectual property or other rights.

Govern 6.2

Contingency processes are in place to handle failures or incidents in third-party data or AI systems deemed to be high-risk.

<u>Map</u>

Map 1

Context is established and understood.

<u>Map 1.1</u>

Intended purpose, potentially beneficial uses, context-specific laws, norms and expectations, and prospective settings in which the AI system will be deployed are understood and documented. Considerations include: specific set or types of users along with their expectations; potential positive and negative impacts of system uses to individuals, communities, organizations, society, and the planet; assumptions and related limitations about AI system purposes; uses and risks across the development or product AI lifecycle; TEVV [test, evaluation, validation, and verification] and system metrics.

<u>Map 1.2</u>

Inter-disciplinary AI actors, competencies, skills and capacities for establishing context reflect demographic diversity and broad domain and user experience expertise, and their participation is documented. Opportunities for interdisciplinary collaboration are prioritized.

<u>Map 1.3</u>

The organization's mission and relevant goals for the AI technology are understood and documented.

<u>Map 1.4</u>

The business value or context of business use has been clearly defined or – in the case of assessing existing AI systems – re-evaluated.

<u>Map 1.5</u>

Organizational risk tolerances are determined and documented.

<u>Map 1.6</u>

System requirements (e.g., "the system shall respect the privacy of its users") are elicited from and understood by relevant AI actors. Design decisions take socio-technical implications into account to address AI risks.

Map 2

Categorization of the AI system is performed.

<u>Map 2.1</u>

The specific task, and methods used to implement the task, that the AI system will support is defined (e.g., classifiers, generative models, recommenders).

Map 2.2

Information about the AI system's knowledge limits and how system output may be utilized and overseen by humans is documented. Documentation provides sufficient information to assist relevant AI actors when making informed decisions and taking subsequent actions.

Map 2.3

Scientific integrity and TEVV considerations are identified and documented, including those related to experimental design, data collection and selection (e.g., availability, representativeness, suitability), system trustworthiness, and construct validation.

Map 3

Al capabilities, targeted usage, goals, and expected benefits and costs compared with appropriate benchmarks are understood.

<u>Map 3.1</u>

Potential benefits of intended AI system functionality and performance are examined and documented.

<u>Map 3.2</u>

Potential costs, including non-monetary costs, which result from expected or realized AI errors or system functionality and trustworthiness – as connected to organizational risk tolerance – are examined and documented.

Map 3.3

Targeted application scope is specified and documented based on the system's capability, established context, and AI system categorization.

<u>Map 3.4</u>

Processes for operator and practitioner proficiency with AI system performance and trustworthiness – and relevant technical standards and certifications – are defined, assessed and documented.

<u>Map 3.5</u>

Processes for human oversight are defined, assessed, and documented in accordance with organizational policies from GOVERN function.

Map 4

Risks and benefits are mapped for all components of the AI system including third-party software and data.

Map 4.1

Approaches for mapping AI technology and legal risks of its components – including the use of thirdparty data or software – are in place, followed, and documented, as are risks of infringement of a third-party's intellectual property or other rights.

<u>Map 4.2</u>

Internal risk controls for components of the AI system including third-party AI technologies are identified and documented.

Map 5

Impacts to individuals, groups, communities, organizations, and society are characterized.

<u>Map 5.1</u>

Likelihood and magnitude of each identified impact (both potentially beneficial and harmful) based on expected use, past uses of AI systems in similar contexts, public incident reports, feedback from those external to the team that developed or deployed the AI system, or other data are identified and documented.

<u>Map 5.2</u>

Practices and personnel for supporting regular engagement with relevant AI actors and integrating feedback about positive, negative, and unanticipated impacts are in place and documented.

<u>Measure</u>

Measure 1

Appropriate methods and metrics are identified and applied.

Measure 1.1

Approaches and metrics for measurement of AI risks enumerated during the Map function are selected for implementation starting with the most significant AI risks. The risks or trustworthiness characteristics that will not – or cannot – be measured are properly documented.

Measure 1.2

Appropriateness of AI metrics and effectiveness of existing controls is regularly assessed and updated including reports of errors and impacts on affected communities.

Measure 1.3

Internal experts who did not serve as front-line developers for the system and/or independent assessors are involved in regular assessments and updates. Domain experts, users, AI actors external to the team that developed or deployed the AI system, and affected communities are consulted in support of assessments as necessary per organizational risk tolerance.

Measure 2

Al systems are evaluated for trustworthy characteristics.

Measure 2.1

Test sets, metrics, and details about the tools used during test, evaluation, validation, and verification (TEVV) are documented.

Measure 2.2

Evaluations involving human subjects meet applicable requirements (including human subject protection) and are representative of the relevant population.

Measure 2.3

Al system performance or assurance criteria are measured qualitatively or quantitatively and demonstrated for conditions similar to deployment setting(s). Measures are documented.

Measure 2.4

The functionality and behavior of the AI system and its components – as identified in the Map function – are monitored when in production.

Measure 2.5

The AI system to be deployed is demonstrated to be valid and reliable. Limitations of the generalizability beyond the conditions under which the technology was developed are documented.

Measure 2.6

Al system is evaluated regularly for safety risks – as identified in the Map function. The Al system to be deployed is demonstrated to be safe, its residual negative risk does not exceed the risk tolerance, and can fail safely, particularly if made to operate beyond its knowledge limits. Safety metrics implicate system reliability and robustness, real-time monitoring, and response times for Al system failures.

Measure 2.7

Al system security and resilience – as identified in the Map function – are evaluated and documented.

Measure 2.8

Risks associated with transparency and accountability – as identified in the Map function – are examined and documented.

Measure 2.9

The AI model is explained, validated, and documented, and AI system output is interpreted within its context – as identified in the Map function – and to inform responsible use and governance.

Measure 2.10

Privacy risk of the AI system – as identified in the Map function – is examined and documented.

Measure 2.11

Fairness and bias – as identified in the Map function – is evaluated and results are documented.

Measure 2.12

Environmental impact and sustainability of AI model training and management activities – as identified in the Map function – are assessed and documented.

Measure 2.13

Effectiveness of the employed TEVV metrics and processes in the Measure function are evaluated and documented.

Measure 3

Mechanisms for tracking identified AI risks over time are in place.

Measure 3.1

Approaches, personnel, and documentation are in place to regularly identify and track existing, unanticipated, and emergent AI risks based on factors such as intended and actual performance in deployed contexts.

Measure 3.2

Risk tracking approaches are considered for settings where AI risks are difficult to assess using currently available measurement techniques or where metrics are not yet available.

Measure 3.3

Feedback processes for end users and impacted communities to report problems and appeal system outcomes are established and integrated into AI system evaluation metrics.

Measure 4

Feedback about efficacy of measurement is gathered and assessed.

Measure 4.1

Measurement approaches for identifying AI risks are connected to deployment context(s) and informed through consultation with domain experts and other end users. Approaches are documented.

Measure 4.2

Measurement results regarding AI system trustworthiness in deployment context(s) and across AI lifecycle are informed by input from domain experts and other relevant AI actors to validate whether the system is performing consistently as intended. Results are documented.

Measure 4.3

Measurable performance improvements or declines based on consultations with relevant AI actors including affected communities, and field data about context-relevant risks and trustworthiness characteristics, are identified and documented.

Manage

Manage 1

Al risks based on assessments and other analytical output from the Map and Measure functions are prioritized, responded to, and managed.

Manage 1.1

A determination is made as to whether the AI system achieves its intended purpose and stated objectives and whether its development or deployment should proceed.

Manage 1.2

Treatment of documented AI risks is prioritized based on impact, likelihood, or available resources or methods.

Manage 1.3

Responses to the AI risks deemed high priority as identified by the Map function, are developed, planned, and documented. Risk response options can include mitigating, transferring, avoiding, or accepting.

Manage 1.4

Negative residual risks (defined as the sum of all unmitigated risks) to both downstream acquirers of AI systems and end users are documented.

Manage 2

Strategies to maximize AI benefits and minimize negative impacts are planned, prepared, implemented, and documented, and informed by input from relevant AI actors.

Manage 2.1

Resources required to manage AI risks are taken into account, along with viable non-AI alternative systems, approaches, or methods – to reduce the magnitude or likelihood of potential impacts.

Manage 2.2

Mechanisms are in place and applied to sustain the value of deployed AI systems.

Manage 2.3

Procedures are followed to respond to and recover from a previously unknown risk when it is identified.

Manage 2.4

Mechanisms are in place and applied, responsibilities are assigned and understood to supersede, disengage, or deactivate AI systems that demonstrate performance or outcomes inconsistent with intended use.

Manage 3

Al risks and benefits from third-party entities are managed.

Manage 3.1

Al risks and benefits from third-party resources are regularly monitored, and risk controls are applied and documented.

Manage 3.2

Pre-trained models which are used for development are monitored as part of AI system regular monitoring and maintenance.

Manage 4

Risk treatments including response and recovery, and communication plans to the identified and measured AI risks are documented and monitored regularly.

Manage 4.1

Post-deployment AI system monitoring plans are implemented, including mechanisms for capturing and evaluating input from users and other relevant AI actors, appeal and override, decommissioning, incident response, recovery, and change management.

Manage 4.2

Measurable activities for continual improvements are integrated into AI system updates and include regular engagement with interested parties, including relevant AI actors.

Manage 4.3

Incidents and errors are communicated to relevant AI actors including affected communities. Processes for tracking, responding to, and recovering from incidents and errors are followed and documented.

NIST AI RMF 1.0, January 2023, Part 2: Core and Profiles, 5. AI RMF Core, https://doi.org/10.6028/NIST.AI.100-1.

Appendix C – Further Reading

Advanced Technology Academic Research Center, From Ethics to Operations: Current Federal Al Policy, January 5, 2022, <u>https://atarc.org/wp-content/uploads/2022/01/Current-Federal-Al-Policy-Assessment-FINAL.pdf</u>.

City of San Jose, Government AI Coalition, accessed March 17, 2024, <u>https://www.sanjoseca.gov/your-government/departments-offices/information-technology/ai-reviews-algorithm-register/govai-coalition</u>.

Congressional Research Service, Artificial Intelligence: Overview, Recent Advances, and Considerations for the 118th Congress, August 4, 2023, <u>https://crsreports.congress.gov/product/pdf/R/R47644</u>.

Daniel Atherton, Reva Schwartz, Peter C. Fontana, Patrick Hall, The Language of Trustworthy AI: An In-Depth Glossary of Terms, March 2023, (National Institute of Standards and Technology, Gaithersburg, MD), NIST. Artificial Intelligence AI 100-3. <u>https://doi.org/10.6028/NIST.AI.100-3</u>.

Executive Order 13960, Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government, <u>https://www.federalregister.gov/documents/2020/12/08/2020-27065/promoting-the-use-of-trustworthy-artificial-intelligence-in-the-federal-government</u>.

Ezzell, Amber, Future of Privacy Forum Generative AI for Organizational Use: Internal Policy Checklist, July 2023, <u>https://fpf.org/wp-content/uploads/2023/07/Generative-AI-Checklist.pdf</u>.

Fennessy, Caitlin, International Association of Privacy Professionals, Regulators' Rulebook for AI: Bit by Bit, August 2, 2023, <u>https://iapp.org/news/a/regulators-rulebook-for-ai-bit-by-bit/</u>.

GSA, AI Guide for Government: A Living and Evolving Guide to the Application of Artificial Intelligence for the U.S. Federal Government, December 3, 2020, <u>https://coe.gsa.gov/coe/ai-guide-for-government/understanding-managing-ai-lifecycle/index.html</u>.

International Association of Privacy Professionals, Global AI Legislation Tracker, August 25, 2023, https://iapp.org/media/pdf/resource_center/global_ai_legislation_tracker.pdf.

International Association of Privacy Professionals, Glossary, n.d., https://iapp.org/resources/glossary/. Microsoft, Governing AI: A Blueprint for the Future, May 2023. https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RW14Gtw.

NIST, U.S. Leadership in AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools, August 9, 2019, https://www.nist.gov/system/files/documents/2019/08/10/ai standards fedengagement plan 9aug2019.pdf.

NIST, AI Risk Management Framework – Resources, Launched March 2023, accessed March 17, 2024, https://www.nist.gov/itl/ai-risk-management-framework/ai-risk-management-framework-resources.

NIST, Trustworthy and Responsible AI Resource Center, Glossary, March 2023, <u>https://airc.nist.gov/AI_RMF_Knowledge_Base/Glossary</u>.

North Carolina Department of Information Technology Enterprise Security and Risk Management Office, Statewide Information Security Manual, January, 2022, <u>https://it.nc.gov/documents/statewide-policies/statewide-information-security-manual/download?attachment</u>

State Archives of North Carolina, Functional Schedule for North Carolina State Agencies Glossary, n.d., https://archives.ncdcr.gov/functional-glossary-key/download?attachment.

White House, Office of Science Technology and Policy, Blueprint for an Al Bill of Rights: Making Automated Systems Work for the American People, n.d., <u>https://www.whitehouse.gov/ostp/ai-bill-of-rights/</u>.