Initial procurement guidelines for public sector procurement, deployment, and monitoring of Generative AI Technology

September 2024

Table of Contents

Introduction

This guidance was created by Washington Technology Solutions (WaTech) in collaboration with the Department of Enterprise Services (DES) and other partners pursuant to the Governor's Executive Order 24-01 to develop initial guidance on the procurement and use of generative artificial intelligence (Gen AI) technology.

Gen AI guidance must address safety and effectiveness, algorithmic discrimination, data privacy, and cybersecurity.

The Executive Order (EO) recognizes that Gen AI is a transformative technology which has the capacity to improve the way the Washington state conducts business and serves the public. The potential to catalyze innovation and enhance human potential is substantial with this new technology. However, this advancement must be tempered with the responsibility to deploy the technology in a safe, responsible, ethical, and efficient manner. As such, the Governor has directed that these Generative AI (Gen AI) guidelines be built on guidance from the White House's Blueprint for an AI Bill of Rights and the National Institute for Science and Technology's (NIST) AI Risk Management Framework and address safety, effectiveness, algorithmic discrimination, data privacy, and cybersecurity. Following the issuance of the Governor's Executive Order, NIST has also released a draft companion resource specific to generative AI which we have also relied on for these guidelines.

"Generative AI Technology" is defined as 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 necessarily identical, to the underlying training data." EO 24-01 defines "High-Risk Generative AI System" and mandates the assessment of high-risk systems prior to deployment. A "High-Risk Generative AI System" is defined as a system "using generative AI technology that creates a high risk to natural persons' health and safety or fundamental rights. Examples include biometric identification, critical infrastructure, employment, health care, law enforcement, and administration of democratic processes." (See Appendix B for levels of risk.)

This document is meant to be a resource for Washington state agencies to use when procuring or using Gen AI. This document provides an overview of principles and steps to take when considering procuring or using Gen AI for public services. This guide does not represent the legal opinion of any Washington state agency, including the Attorney General's Office. Readers should not rely on information in this guide regarding specific applications of the laws without seeking legal counsel.

Relevant legal and regulatory requirements

All Generative AI procurement or use guidance shall defer to and work in concert with procurement rules and laws in Washington including but not limited to Chapter 39.26 RCW, Washington Department of Enterprise Services procurement policies and rules, and WaTech technical and security policies.

Deployment considerations

This document includes a series of recommendations that should be considered when evaluating the procurement, development, deployment, use, or ongoing monitoring of Generative AI technology. These guidelines include tasks that should take place throughout the lifecycle of use of a Gen AI technology. For purposes of this procurement guidance, the recommendations are organized into three different phases:

1. Requirements development.
2. Procurement, development , and deployment.
3. Ongoing monitoring.

The purpose of this document is to provide procurement guidance that aligns with the:

- White House's Blueprint for an AI Bill of Rights.
- National Institute for Science and Technology's AI Risk Management Framework.
- Washington state's interim guidelines for purposeful and responsible use of artificial intelligence.

This guidance includes tasks and considerations of risks, as well as actions or controls agencies can undertake to address risks and align with these underlying principles.

In determining whether to proceed with the procurement, development, deployment, or use of Gen AI, agencies should exercise discretion and thoughtfulness. In doing so, agencies should consider the risks and goals of using Gen AI, which are described in the Core Concepts section. With these considerations in mind, agencies should consider the appropriateness of progressing with the procurement or use of Gen AI and the business problem its use seeks to address. As such, agencies should keep in mind that:

- The capabilities of Gen AI technology may change over time with new productions or version releases.
- Changes in Gen AI technology capabilities may change associated risks for both the agency and Washington residents, which could change review or approval requirements for development, procurement, or usage of particular Gen AI technology.
- WaTech may adjust and update this Gen AI procurement, development, deployment, use, and ongoing monitoring guidance based on updated government and industry research, experiences, or legal requirements.
- Agencies should be ready to be intentional and descriptive in their assessments and are encouraged to seek support and guidance from WaTech.

Core concepts

Definition of Generative AI Technology

For purposes of this initial Gen AI Procurement, Use, and Monitoring Guidance, the following definitions of "Generative AI Technology" and "High-Risk Generative AI System" shall be used (See EO 24-01):

- "Generative AI Technology" 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.
- "High-Risk Generative AI System" means systems using generative AI technology that creates a high risk to natural persons' health and safety or fundamental rights. Examples include biometric identification, critical infrastructure, employment, health care, law enforcement, and administration of democratic processes.

Goals of using Generative AI Technology

There are several goals of using generative AI technology, including:

- Efficiency: Gen AI may assist governmental entities in streamlining processes and making administrative decisions more quickly and efficiently.

- Economical: Gen AI may reduce costs by reducing the time necessary to record and analyze data, conduct calculations, and make decisions.

- Equity: Reduce the biases and inaccuracies of current systems: While Gen AI may introduce and reinforce bias, it also potentially provides an opportunity to address biases.

- Effectiveness: Improving delivery of public services: Gen AI may help the public receive improved and more accessible services.

Risks

Conversely, there are several known risks of using Gen AI technology today:

- Inaccuracies: Gen AI technologies may reproduce biases or introduce new inaccuracies or "hallucinations" and may be less accurate than human decision-makers.

- Automation bias: Humans tend to place too much trust in automated decisions. Humans may overestimate the accuracy of data analysis outputs by Gen AI systems, which may be as, or more, error-prone than human decision-makers. Automation bias can obfuscate or increase biased and inaccurate decision-making. This may include discounting contradictory information made without automation.

- Non-transparency: It is difficult or impossible for individuals to know if an Gen AI is being used, how that system operates, and the impacts of the system on individuals and society.

- Lack of explainability: Gen AI outputs are very difficult to explain in clear and concise language that would be understandable to those auditing the system or those potentially impacted by their use. This risk may be especially prevalent when a Gen AI tool is procured through a third-party vendor. Even the developers of Gen AI models may not know why or how particular outputs are provided due to the size and complexity of large models.

- Lack of accountability: Individuals who are affected by Gen AI outputs may not have the ability to meaningfully challenge a system's decisions. Governmental entities that procure or use Gen AI systems may overlook consulting with the individuals and communities that may be affected by its use or may not have a human-centered dispute resolution process.

- Threats to privacy: Large amounts of data about individuals are often used to train Gen AI technologies (both simple and complex) to transform inputs into decisions or suggestions. Individuals may not understand or have given consent for their data to be used for such a purpose.

- Threats to legitimacy and public trust: Use of Gen AI systems may undermine the legitimacy and public trust of governmental entities when such entities re-delegate their decision-making responsibilities to unaccountable and nontransparent systems.

- Discrimination: Gen AI outputs may reproduce or exacerbate existing patterns of discrimination that are already present in our society. This can lead to harms on vulnerable or marginalized communities.

- Threats to cybersecurity: Cybersecurity risks include prompt injections, malicious code, accidental data leaks and breaches due to control or permission issues when using unproperly reviewed resources or systems.

Washington State Agency Privacy Principles

The Washington State Agency Privacy Principles serve as a framework to guide state agencies in protecting the privacy and security of personal information they collect and maintain. These principles were developed to address the growing concerns surrounding data privacy and ensure that individuals' personal information is handled with care and transparency. Applying these principles helps agencies balance the risks and goals of systems processing personal information, especially in high-risk Generative AI systems. Along those lines, these principles promote accountability, transparency, and the responsible use of personal data by state agencies.

1. **Lawful, fair, and responsible use:** The collection, use and disclosure of information is based on legal authority, not deceptive, not discriminatory or harmful, and relevant and reasonably necessary for legitimate purposes.

2. **Data minimization:** The minimum amount of information is collected, used, or disclosed to accomplish the stated purpose for collecting the information.

3. **Purpose limitation:** The reasons for gathering information are identified before it is collected. Use and disclosure is limited to what is reasonably necessary in relation to the specific reasons the information was collected.

4. **Transparency and accountability:** Transparency means being open and transparent about what personal information is collected, for what purposes, and who it is shared with under what circumstances. Accountability means being responsible and answerable for following data privacy laws and principles.

5. **Due diligence:** Taking reasonable steps and exercising care before and after entering into an agreement or arrangement with a third party that includes sharing personal information.

6. **Individual participation:** Give people control of their information when possible.

7. **Security:** Appropriate administrative, technical and physical security practices to protect the confidentiality, integrity, availability and control of personal information.

The Washington State Agency Privacy Principles aim to establish a culture of privacy and data protection within state agencies, promoting trust and confidence among individuals whose personal information is collected. By adhering to these principles, state agencies can ensure the responsible handling of personal data while balancing the need for data-driven decision-making and public service delivery which may use Gen AI.

Generative AI Technology Guiding Principles

The intention of the state of Washington is to follow the principles in the White House Blueprint for an AI Bill of Rights and the National Institute of Standard and Technology (NIST) AI Risk Framework, which serve as the basis for the purposeful and responsible use of artificial intelligence. A foundational part of both the Blueprint and NIST AI Risk Framework is to ensure the trustworthiness of systems that use AI, which includes Generative AI. The guiding principles are:

- **Safe, secure, and resilient:** AI should be used with safety and security in mind, minimizing potential harm and ensuring that systems are reliable, resilient, and controllable by humans. AI 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 non-discrimination:** AI must be developed and used 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:** AI 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 design, development, and deployment. Privacy-enhancing systems should safeguard human autonomy and identity where appropriate.

- **Accountability and responsibility:** As public stewards, agencies should use AI responsibly and be held accountable for the performance, impact, and consequences of its use in agency work.

- **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 a AI system and its outputs are available to the individuals interacting with the system. Transparency answers "what happened" in the system.

- **Explainable and interpretable:** Agencies should ensure AI technology use in a system can be explained, meaning "how" the decisions or outputs that were 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. Where appropriate, individuals should be able to opt out of Gen AI technology or high-risk Generative AI systems in favor of a human alternative.

Generative AI procurement guidance

This guidance is for Generative AI solutions that are both intentional and incidental to technology procurements. Intentional generative AI is where an agency is specifically seeking to purchase or use a generative AI product or service. Incidental generative AI is when the Gen AI capability is built into the product or service, which may include software as a service (SaaS) solutions. However, please note this document is intended to be best practices and recommendations for how to think about procuring and deploying Generative AI responsibly.

Phases

1. Requirements Development.
2. Procurement, development, and deployment.
3. Ongoing Monitoring.

Requirements development

The Requirements Development phase is the initial stage of the process where an assessment is conducted before the system's development or procurement. It involves evaluating the reasons for adopting the system and identifying the associated risks in its design and implementation. This phase focuses on considering factors such as:

- The system's benefits in fulfilling the agency's mission and societal goals.
- Identifying risks for potential inaccuracies, biases, or disproportionate effects.
- Considering the system's needed security against data disclosure or manipulation.
- Determining type and amount of data needed.
- Assessing its potential impact on public trust.

The phase also emphasizes the importance of providing opportunities for public participation, documenting the process and decisions transparently, and exercising caution in adopting new systems that lack testing for bias or accuracy. This phase sets the foundation for informed decision-making and paves the way for subsequent development or procurement stages.

Procurement, development, and deployment

The Procurement, Development, and Deployment phase refers to the stage in which the system is either built or acquired for deployment. During this phase, careful consideration is given to the system's design, functionality, and the procurement process involved. It entails translating the identified requirements and objectives into a tangible system, ensuring that it aligns with the intended purpose and adheres to applicable guidelines and regulations. This phase involves activities such as system development, vendor selection, contract negotiation, and the establishment of technical specifications and deployment plans. It also includes conducting thorough testing and validation to assess the system's performance, accuracy, and potential biases before its deployment.

This procurement phase is critical in ensuring that the Gen AI technology is properly designed, sourced, and prepared for subsequent operational and deployment phases. These guidelines require that any vendor providing a high risk Generative AI system to an agency must certify that the vendor has implemented an AI governance program consistent with the National Institute of Standards and Technology Artificial Intelligence Risk Management Framework. (See Executive Order 24-01)

Ongoing monitoring and updated testing and assessments should be required when substantive system modifications or changes in the Gen AI use occur.

Ongoing monitoring

The Ongoing Monitoring phase refers to the continual process of observing, evaluating, and assessing the system's performance, outcomes, and adherence to established standards. It involves the regular monitoring of data inputs, system outputs, and decision-making processes to identify any potential issues, biases, inaccuracies, or unintended consequences. This phase includes the collection and analysis of relevant data, such as audit trails, logs, or hashes to evaluate the system's effectiveness, fairness, and reliability over time.

Ongoing monitoring also involves periodically comparing the system's outcomes against desired goals and benchmarks to ensure its alignment with organizational objectives and societal values. Additionally, it may encompass periodic assessments, independent audits, and participant engagement to promote transparency, accountability, and ongoing improvement of the Gen

AI technology. The monitoring phase aims to maintain the system's integrity, address emerging challenges, and facilitate necessary adjustments and enhancements to optimize its performance and mitigate potential risks. Ongoing monitoring and updated testing and assessments should be required when substantive system modifications or changes in the Gen AI use occur.

Generative AI procurement tasks

Ten activities should be completed during the procurement or deployment of a Gen AI technology, which includes both intentional Gen AI and incidental Gen AI. "Generative AI Technology" is defined as 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 necessarily identical, to the underlying training data." (EO 24-01).

As shown in the chart below, some of these activities are tied to a specific phase, while others span two or three phases. These tasks should be considered throughout the Gen AI technology use but are not exhaustive or exclusionary of other considerations. Detailed descriptions of the activities and related tasks are below. The activities include:

1. Gen AI initial use case assessment.
2. Evaluation of whether to adopt a system.
3. Updated assessments.
4. Periodic testing.
5. Transparency of the model or system.
6. Audit trails.
7. Training to understand automation basis.
8. Evaluation of risks/determination whether to proceed.
9. Review of decisions by those affected.
10. Weighing advantages against known bias or inaccuracies.

Generative AI procurement tasks table

	Phases	Requirements development	Procurement, development, and deployment	Ongoing monitoring
1.	Gen AI Initial Use Case Assessment.	√	√	
2.	Evaluation of whether to adopt a system.	√		
3.	Updated assessments.			√
4.	Periodic testing.			√
5.	Transparency of the model or system.	√	√	√
6.	Audit trails.	√	√	√
7.	Training to understand automation basis.	√	√	√
8.	Evaluation of risks/determination whether to proceed.	√	√	√
9.	Review of decisions by those affected.	√	√	√
10.	Weighing advantages against known bias or inaccuracies.		√	√

Purpose of Generative AI tasks

The purpose of the ten tasks for Gen AI procurement, development, deployment, or use is to mitigate the risk and help achieve the goals of implementing Gen AI technology. This generally includes practices such as:

- An evaluation of whether it is appropriate for governmental entities to use the Gen AI technology. Depending on system impacts, this evaluation should incorporate public participation and comment.
- Independent testing for bias and inaccuracy.
- Transparency so that the public is informed that a Gen AI technology is being used and understands information about the system and its use.
- Accountability so that the public may meaningfully engage and allow individuals to seek redress when impacted by a Gen AI technology.

Description of tasks

This section provides a description of the Gen AI technology procurement tasks identified above and the phase of the Gen AI procurement lifecycle each activity should be performed.

- **Generative AI Initial Use Case Assessment**: Prior to procurement, government organizations need to determine what is the business need or problem they are trying to solve. While Gen AI may be an appropriate solution, it may not always be optimal or necessary for an agency's needs. Agencies should perform an Initial Generative AI Use Case Assessment for either intentional or incidental Gen AI (See example assessment in Appendix B) that informs the decision to pursue a Gen AI technology solution. This assessment includes considering cost-benefit analysis and operational benefits of procuring, developing, or deploying a Gen AI tool. As part of this process, agencies need to also determine if the use case meets the definition of "high-risk."

  The Initial Gen AI Use Case Assessment considers various criteria to determine resource investment. These criteria may include factors such as:
  - The significant impact on identified or identifiable individuals.
  - The scale of influence on many individuals.
  - The potential for high error risks (especially for systems lacking transparency, bias testing, or accuracy assessment).
  - The agency's discretion in creating the model or system.
  - The level of automation with opportunities for human review.
  - How the Gen AI model will be used within a system.

When the Gen AI technology significantly affects individuals in high-risk use cases, the governmental entity should be transparent about its initial assessment of the Gen AI technology and make it available to the public.

Phase(s):

- Requirements development.
- Procurement, development, and deployment.
- Ongoing monitoring.

- **Evaluation of whether to adopt a system**: Before acquiring, creating, or employing a Gen AI technology, the government organization must conduct an evaluation of the rationale behind adopting such systems, identify business needs, and the associated risks related to their design and deployment. This assessment should leverage and optimize the existing procurement process to ensure the most effective use of resources. The evaluation should consider the following:

    - Assess how the system contributes to the agency's mission and societal objectives.
    - Conduct tests to identify any potential inaccuracies, biases, or disproportionate effects arising from the system or the data sources used in its design or training. If such issues are recognized, identify appropriate measures that should be taken to address them.
    - Ensure the Gen AI system's resilience against unauthorized data disclosure or malicious manipulation (e.g. data poisoning, deep fakes, or unethical or offensive misuse).
    - Determine whether the use of a Gen AI may adversely impact the public's trust in the actions of the Washington state government.
    - Provide opportunities for public participation, including informing the public about the risks and benefits associated with the Gen AI technology and soliciting meaningful input from affected individuals and communities regarding its design and deployment.
    - Document the process and decisions for adoption in a manner that enables future review while maintaining transparency for the public.
    - Exercise extreme caution when considering the adoption of any new high-risk Gen AI system that will impact individuals' rights and freedoms, employment, health care, or other democratic or judicial process that has not undergone bias or accuracy testing. And ensure that the testing process includes appropriate disclosures to facilitate

independent review by members of the public or independent entities.

By thoroughly evaluating these factors, the governmental entity can make informed decisions regarding the acquisition, development, and usage of Gen AI technology, prioritizing transparency, accountability, and public trust.

Phase:

- Requirements development.

- **Updated assessments**: The evaluation of deployed Gen AI technology should be periodically reviewed and re-evaluated whenever there are subsequent modifications made to the Gen AI itself or to the data collection process (including new training data) that informs the model. These updates and reassessments must be conducted in a transparent manner, ensuring that the public is kept informed of any changes.

  Phase:

  - Ongoing monitoring.

- **Periodic testing**: It is important to establish a process to regularly test the performance of the Gen AI while it is in use. This periodic testing aims to identify any indications of inaccuracies, biases, or unequal outcomes that may arise from the system, and should be conducted at intervals in proportion to the potential for risk. If such tendencies are detected, they should be promptly addressed, or if there are reasons for not addressing them, those reasons must be clearly explained in detail. The entire process and the subsequent testing results should be transparent and made available to the public. The establishment of a regular testing process, testing, and transparency of results ensures the fairness and accountability of the system's implementation by the government. For High-Risk Gen AI systems an independent third-party audit is highly recommended for periodic testing to ensure impartial review.

  Phase:

  - Ongoing monitoring.

- **Transparency of training data, inputs or outputs**: While not always possible, it is encouraged that the training data, inputs, and outputs of the Gen AI systems be well understood by the agency. However, certain compelling reasons may exist to restrict the release, such as cases where

there is a clear and demonstrable threat to governmental integrity or a significant risk that individuals may exploit the system, leading to a threat to governmental integrity or a significant cybersecurity risk. In situations where commercial or government interests justify the restriction of model training data, third-party evaluations should be conducted to assess the potential risks of inaccuracy or bias. It's important to note that while reviewing the training data is important, it cannot serve as a substitute for comprehensive testing to identify any inaccuracies or biases in the design and implementation of the Gen AI.

Phase(s):

- Requirements development.
- Procurement, development, and deployment.
- Ongoing monitoring.

- **Audit trails**: The Gen AI should produce audit trails, logs, or hashes that document the inputs and outputs in its Gen AI process that are contained in system logs. This enables the governmental entity to furnish individuals with the rationale behind the Gen AI's determinations and simplifies potential future scrutiny of those outputs. The recorded information should also be accessible to third-party researchers, granting them the opportunity to conduct impartial assessments of the system's accuracy and potential biases.

  Phase(s):

  - Requirements development.
  - Procurement, development, and deployment.
  - Ongoing monitoring requirements development.

- **Training to understand automation bias**: Individuals engaged in the procurement, development, or operation of Gen AI should undergo training that explicitly addresses the concept of automation bias. Automation bias is an over-reliance on automated aids and decision support systems. As the availability of Gen AI outputs increases, it is a human tendency to rely on system outputs without question. Training on automation bias makes the workforce aware of these tendencies to mitigate against harmful or biased system outputs.

  Phase(s):

  - Requirements development.
  - Procurement, development, and deployment.
  - Ongoing monitoring.

- **Evaluation of risks/determination whether to proceed**: Before deploying the Gen AI, and whenever there is a discovery of potential inaccuracies or biases, the governmental entity must assess whether the risks and impacts of such inaccuracies or biases on individuals and the potential erosion of public trust are significant enough to warrant not using the system. In making this determination, input from the individuals who will be impacted by the system should be considered along with other risks or harmful impacts that could emerge. The final decision should be clearly outlined in written form, transparent, and made available to the public.

  Phase(s):
  - Requirements development.
  - Procurement, development, and deployment.
  - Ongoing monitoring.

- **Review of decisions by those affected**: Individuals impacted by a decision made or facilitated by a High-Risk Gen AI system should have the ability to examine and question the underlying foundation of that decision, especially in cases where the impacts to rights or responsibilities of individuals are significant. Part of the requirements development phase and development and procurement phase should include an evaluation of whether the Gen AI will be high risk and impact the rights and freedoms of individuals. If so, the ability to perform decision or output review should be part of those procurement phases as well as the ongoing monitoring phase for actual review. People should be made aware of any usage of a Gen AI system that affects them, should have the opportunity to opt-in or opt-out of the use of the Gen AI system, and should have clear channels through which to report issues or complaints.

  Phase(s):
  - Requirements development.
  - Procurement, development, and deployment.
  - Ongoing monitoring.

- **Weighing advantages against known bias or inaccuracies**: While acknowledging the significance of the benefits generated by Gen AI technology, it is essential to prioritize the prevention of harm to individuals caused by known or potential biases or inaccuracies. When weighing the advantages of benefits against the potential harm to natural persons, the avoidance of harm should be assigned considerably greater importance.

  Government organizations should assess the extent to which the acquisition and deployment of their Gen AI technology adheres to the Gen AI Guiding Principles and Washington State Privacy Principles. If it is discovered that the procurement, use, or deployment of a system fails to align with these principles, the governmental entity should identify the reasons for such non-alignment and take suitable actions to mitigate its uses based on the guiding principles.

  Phase(s):
  - Procurement, development, and deployment.
  - Ongoing monitoring.

Actions or controls to take for tasks

Each table below correlates to an activity described above. For each activity there are actions or controls that agencies can take to mitigate risks and align to the Washington State Agency Privacy Principles and AI Guiding Principles for purposeful and responsible use of Gen AI technology. There may be additional risks and principles addressed by controls agencies choose to undertake with tasks, but the tables help identify predominant risks and principles.

As Gen AI technologies are used, state agency staff should continue to work with WaTech's technology and privacy staff regarding implicated risk for the state and its residents. Agency staff should complete and update its Gen AI inventory to maintain a comprehensive record of Gen AI used throughout the state. Appendix A lists additional sample questions to identify and assess the Gen AI and its use to the agency in performing work for Washington.

1. GENERATIVE AI INITIAL USE CASE ASSESSMENT FOR INTENTIONAL OR INCIDENTAL GEN AI			
Action or Control	**Notes**	**Risk Addressed**	**Principles**
What is the business need or problem your agency is trying to solve?	Has the agency defined a clear purpose addressable by Gen AI technology? Is it the most appropriate tool for the solution?	Lack of accountability	Public purpose and social benefit
Is this a high-risk use case?	Consider risks to rights and freedoms to individuals. Is this a high-risk Gen AI system with potential impacts to critical infrastructure? Health? Safety? Employment? Democratic or judicial process? Biometric identification or verification? Surveillance?	Discrimination. Lack of explainability. Threats to privacy.	Lawful, fair and responsible use. Fairness, inclusion, and non-discrimination. Explainable and interpretable. Public purpose and social benefit.
Consider the cost-benefit analysis of procuring, developing, or using a Gen AI system.		Lack of accountability.	Due diligence. Accountability & responsibility.
Consider operational benefits of procuring, developing, or using a Gen AI system.		Lack of accountability.	Due diligence. Accountability & responsibility.
Use Gen AI Initial Use Case Assessment and Risk/Impact Matrix.		Non-transparency. Lack of accountability.	Purpose limitation. Accountability & responsibility.

2. EVALUATION OF WHETHER TO ADOPT SYSTEM			
Action or Control	**Notes**	**Risk Addressed**	**Principles**
Consider whether Gen AI is able to address the identified problem or issue.	Has the agency defined a clear purpose in using the identified Gen AI system?	Non-transparency. Lack of accountability. Threats to legitimacy and public trust.	Lawful, fair and responsible Use Purpose limitation. Transparency & accountability. Public purpose & social benefit. Safe, secure, resilient.Fairness, inclusion, and non-discrimination.
Consider whether there is agency leadership support for utilization of the Gen AI technology.	Has agency leadership defined clear goals and success criteria for adoption of the Gen AI technology?	Non-transparency. Lack of accountability. Threats to legitimacy and public trust.	Lawful, fair and responsible use. Purpose limitation. Transparency & accountability. Public purpose & social benefit. Safe, secure, resilient. Fairness, inclusion, and non-discrimination.
	Have responsibilities of agency staff involved in adoption of Gen AI technology been delineated and communicated to: - Business staff. - Technical staff. - Executive/strategic level staff.	Non-transparency. Lack of accountability. Threats to legitimacy and public trust.	Lawful, fair and responsible use. Purpose limitation. Transparency & accountability. Public purpose & social benefit.
Consider whether there is agency leadership support for utilization of the Gen AI technology.	Differentiate between technical and data protection risks.	Discrimination. Inaccuracies. Lack of accountability. Threats to privacy.	Lawful, fair and responsible use. Purpose limitation. Transparency & accountability. Fairness, inclusion, and non-discrimination. Privacy and data protection. Transparency & auditability.

2. EVALUATION OF WHETHER TO ADOPT SYSTEM			
Action or Control	**Action or Control**	**Action or Control**	**Action or Control**
Consider the level of agency human worker involvement in Gen AI technology used for decision-making.	Human worker involvement/interaction: - Human worker-in-the-loop (human decision-making supports the process). - Human worker-out-of-the-loop (human decision-making is not part of the process). - Human worker-over-the-loop (human intervention may occur in the decision-making process).	Human-based errors. Discrimination. Inaccuracies. Non-transparency. Lack of explainability.	Transparency & accountability. Valid and reliable. Safe, secure, resilient. Fairness, inclusion, and non-discrimination.
	Factors to consider: - Risk tolerance – what is the risk tolerance of the agency for human worker involvement/interaction with Gen AI technology. - Individual user-experience – is the user-experience enhanced or reduced by the involvement/interaction by a human worker in the Gen AI technology user flow. - Operational cost – what is the cost-benefit of a human worker involving/interacting with a Gen AI technology.	Human-based errors. Discrimination. Inaccuracies. Non-transparency. Lack of explainability.	Transparency & accountability. Explainability & interpretable. Accountability & responsibility.

3. UPDATED ASSESSMENTS			
Action or Control	**Notes**	**Risk Addressed**	**Principles**
Assess or re-evaluate Gen AI if significant modifications are made to the model or system.	Factors to consider: - New training data - Changes in model design - New use case(s)	Non-transparency. Lack of accountability. Threats to legitimacy and public trust.	Lawful, fair and responsible use. Purpose limitation. Transparency & accountability. Transparency & auditability. Accountable & responsible.
Assess or re-evaluate Gen AI if subsequent changes are made to data collection process.		Non-transparency. Lack of accountability. Threats to legitimacy and public trust.	Lawful, fair and responsible use. Purpose limitation. Transparency & accountability. Privacy and data protection. Accountable & responsible.
Assess or re-evaluate Gen AI if significant changes are made to training data that informs the model.	Factors to consider: - New training data - Changes in model design - New use case(s)	Non-transparency. Lack of accountability. Threats to legitimacy and public trust.	Lawful, fair and responsible use. Purpose limitation. Transparency & accountability. Valid and reliable. Accountability and responsibility.

4. PERIODIC TESTING			
Action or Control	**Notes**	**Risk Addressed**	**Principles**
Develop evaluation and testing plan for Gen AI technology.	Testing should be proportionate to potential risks of the Gen AI technology.	Lack of accountability. Non-transparency. Threats to legitimacy and public trust.	Lawful, fair and responsible use. Accountability and responsibility.
Create a process for periodic testing.	Process to include how users or public can flag issues related to bias, discrimination, or poor performance.	Automation bias. Inaccuracies. Discrimination.	Lawful, fair, and responsible use. Privacy and data protection. Valid and reliable.
Implement a plan and use independent third-party to evaluate risk.	Most likely to use independent third parties for High-Risk Gen AI systems.	Automation bias. Inaccuracies. Discrimination.	Due diligence. Valid and reliable. Privacy and data protection.
Publish results.		Non-transparency. Threats to legitimacy and public trust.	Transparency & accountability. Transparency & auditability. Accountability and responsibility.

5. TRANSPARENCY OF ALGORITHM, TRAINING DATA, INPUTS, or OUTPUTS			
Action or Control	**Notes**	**Risk Addressed**	**Principles**
Consider whether the Gen AI has a data model developed or if the agency needs to develop a data model with agency data.	If the agency needs to develop a data model for evaluation, development and/or procurement of the Gen AI system, utilize the agency designated data scientist and/or liaise with the WaTech office for guidance.	Discrimination. Inaccuracies. Automation bias. Non-transparency. Lack of explainability.	Data minimization. Purpose limitation. Transparency & accountability. Transparency & auditability. Valid and reliable. Explainable & interpretable.
Consider whether the Gen AI has a data model developed or if the agency needs to develop a data model with agency data.	If the data model exists, does it conform to the Washington State Agency Privacy Principles and the Gen AI Guiding Principles?	Threat to privacy.	Data minimization. Purpose limitation. Transparency & accountability. Privacy and data protection.
Consider whether the Gen AI has a data model developed or if the agency needs to develop a data model with agency data.	If individual or client data is used to develop data model, did the agency review the responses to minimize false, misleading or inaccurate responses?	Discrimination. Inaccuracies. Automation bias. Non-transparency. Lack of explainability.	Data minimization. Purpose limitation. Transparency & accountability. Privacy and data protection. Safe, secure, and resilient.

5. TRANSPARENCY OF ALGORITHM, TRAINING DATA, INPUTS, or OUTPUTS			
Action or Control	**Action or Control**	**Action or Control**	**Action or Control**
Consider relevant features or functionalities that have greatest impact on the agency client or customers affected by the Gen AI technology.	Have features or functionalities that will improve trust with the agency been identified?	Threats to legitimacy and public trust.	Lawful, fair and responsible use. Data minimization. Purpose limitation. Accountability & responsibility.
Consider relevant features or functionalities that have greatest impact on the agency clients or customers affected by the Gen AI technology. (cont.)	Have features or functionalities that will reduce trust/harm reputation of the agency been identified?	Threats to legitimacy and public trust.	Lawful, fair and responsible use. Data minimization. Purpose limitation. Fairness, inclusion, and non-discrimination.
Has the algorithm, inputs and outputs of the Gen AI been assessed as valid by an independent third-party?		Inaccuracies. Non-transparency. Lack of explainability. Lack of accountability.	Transparency & accountability. Valid and reliable. Transparency & auditability.
Can the agency describe in "plain English" how the production version of the Gen AI technology makes decisions or generates outputs?	Can reports be generated by providing detail on the explainability of Gen AI technology features and functionalities.	Non-transparency. Lack of explainability. Lack of accountability.	Transparency & accountability. Individual participation. Explainable & interpretable.

5. TRANSPARENCY OF ALGORITHM, TRAINING DATA, INPUTS, or OUTPUTS			
Action or Control	**Action or Control**	**Action or Control**	**Action or Control**
	Consider formally documenting and publishing an FAQ about the Gen AI technology, including operational details of the generating or decision-making model.	Non-transparency. Lack of explainability. Lack of accountability. Threats to legitimacy and public trust.	Transparency & accountability. Individual participation. Explainable & interpretable.
Can the agency describe in "plain English" how the production version of the Gen AI technology makes decisions or generates outputs? (cont.)	Document potential limitations and gaps of the Gen AI technology generating or decision-making model.	Discrimination. Inaccuracies. Automation bias. Non-transparency. Lack of explainability.	Transparency & accountability. Accountable and responsible. Transparency & auditability.

5. TRANSPARENCY OF ALGORITHM, TRAINING DATA, INPUTS, or OUTPUTS			
Action or Control	**Action or Control**	**Action or Control**	**Action or Control**
If certain features or functions of the Gen AI technology or model cannot be explained in "plain English," consider conducting additional tests in pre-production environments and additional user-acceptance tests with clients or customers.	Hold forums for public or customer feedback.	Lack of accountability. Threats to legitimacy and public trust.	Transparency & Accountability. Individual participation. Transparency & auditability. Fairness, inclusion, and non-discrimination.
	Work with technical staff to define different models of production implementation for testing against various customer subpopulations against defined success criteria.	Human-based errors. Discrimination. Inaccuracies. Automation bias. Lack of explainability.	Transparency & accountability. Fairness, inclusion, and non-discrimination. Valid and reliable.
Consider updating agency privacy policy with information regarding use of Gen AI systems, including information collected and processed, and how decisions are made.	Link to a developed FAQ.	Non-transparency. Lack of accountability. Threats to privacy.	Transparency & Accountability. Individual participation. Transparency & auditability. Explainable and interpretable.

5. TRANSPARENCY OF ALGORITHM, TRAINING DATA, INPUTS, or OUTPUTS			
Action or Control	**Action or Control**	**Action or Control**	**Action or Control**
	Consider providing a just-in-time notice for individuals regarding the use of Gen AI technology (e.g. chatbots).	Non-transparency. Lack of accountability. Threats to privacy.	Transparency & accountability. Individual participation. Accountable and responsible.
Consider providing notice of intent for use of High-Risk Gen AI technology to public before procurement, deployment, or use.	See outline of process in RCW 43.386 re Facial Recognition for example.	Non-transparency. Lack of accountability. Threats to privacy	Transparency & accountability. Individual participation. Accountable and responsible.
Consider updating data sharing agreements (DSAs) if Gen AI systems are used in the business-to-business relationship.	Refer to DSA implementation guidance.	Non-transparency. Lack of accountability. Threats to privacy. Threats to legitimacy and public trust.	Transparency & accountability. Due diligence. Privacy and data protection.

6. AUDIT TRAILS			
Action or Control	**Notes**	**Risk Addressed**	**Principles**
Consider establishing a knowledge management repository/database to record/archive relevant documents regarding developing, procuring, or using Gen AI technology.	Consider documenting factors in making decision to develop, procure, or use a Gen AI system.	Non-transparency. Lack of accountability. Threats to legitimacy and public trust.	Due diligence. Transparency & auditability.
Formally document change management program for Gen AI technology.	Consider setting cadence of change management meetings for Gen AI technology.	Inaccuracies. Non-transparency. Lack of explainability. Lack of accountability.	Transparency & accountability. Security. Safe, secure, and resilient. Accountability and responsibility.
	Include assessment of implicated privacy risk per changes in change management program. Coordinate with WaTech.	Inaccuracies. Non-transparency. Lack of explainability. Lack of accountability.	Transparency & accountability. Security. Accountability and responsibility. Privacy and data protection.

6. AUDIT TRAILS			
Action or Control	**Notes**	**Risk Addressed**	**Principles**
Establish tracking and monitoring of role-based privacy-relevant training on automation bias and additional relevant data protection implications of using Gen AI.		Discrimination. Automation. Bias. Lack of explainability. Threats to privacy. Threats to legitimacy and public trust.	Lawful, fair and responsible use. Transparency & accountability. Fairness, inclusion, and non-discrimination. Accountability and responsibility.
If the Gen AI technology fails or flags a failure or discrepancy in the decision/projected answer, is there a pre-determined threshold? Consider how staff workers will be notified of this prior to a decision being made by the high risk Gen AI system.	Consider implications to business continuity plan.	Inaccuracies. Automation. Bias. Lack of explainability.	Lawful, fair and responsible use. Transparency & accountability. Security. Valid and reliable.
Consider developing formalized feedback loop with affected clients or customers.	Use feedback from clients or customers early in the development & procurement cycle to support identification of success criteria.	Lack of accountability. Threats to legitimacy and public trust.	Individual participation. Public purpose and social benefit. Fairness, inclusion, and non-discrimination. Accountability and responsibility.

6. AUDIT TRAILS			
Action or Control	**Action or Control**	**Action or Control**	**Action or Control**
	Review risk reduction/ acceptance of feedback throughout development & procurement cycle.	Lack of accountability. Threats to legitimacy and public trust.	Individual participation. Accountability and responsibility.
	Document closed feedback loop; document report for publication.	Lack of accountability. Threats to legitimacy and public trust.	Individual participation. Fairness, inclusion, and non-discrimination. Accountability and responsibility.

7.TRAINING TO UNDERSTAND AUTOMATION BIAS			
Action or Control	**Notes**	**Risk Addressed**	**Principles**
Consider whether agency staff involved in the development, procurement, or use of Gen AI technology have been trained on automation bias.		Discrimination. Inaccuracies. Automation bias. Lack of accountability.	Lawful, fair and responsible use. Purpose limitation. Transparency & accountability. Fairness, inclusion, and non-discrimination. Accountability and responsibility.
Consider whether agency staff involved in the development, procurement, or use of Gen AI technologies have training on Washington State Agency Privacy Principles and responsible use of Gen AI.		Lack of accountability. Threats to privacy.	Lawful, fair and responsible use. Purpose limitation. Transparency & accountability. Privacy and data protection.
Consider designating a specialized role with specific responsibilities regarding ethical and data protection issues of utilizing high risk Gen AI systems within agency.	Alternatively, consider formalizing liaison relationship with WaTech for such guidance.	Human-based errors. Inaccuracies. Lack of explainability.	Lawful, fair and responsible use. Purpose limitation. Transparency & accountability. Fairness, inclusion, and non-discrimination. Accountability and responsibility.

8. EVALUATION OF RISKS/DETERMINATION WHETHER TO PROCEED			
Action or Control	**Notes**	**Risk Addressed**	**Principles**
Consider whether agency has ability to test Gen AI systems before procurement or deployment.	Is there a sand-box, development, testing or QA environment that can be used to test implementation/integration of Gen AI technology prior to production.	Inaccuracies. Non-transparency. Lack of explainability.	Lawful, fair and responsible use. Transparency & accountability. Safe, secure, and resilient. Valid and reliable.
	Is there an agreed-upon requirements and/or punch list that the Gen AI technology should satisfy before moving to production?	Inaccuracies. Non-transparency. Lack of explainability.	Lawful, fair and responsible use. Transparency & accountability. Safe, secure, and resilient. Valid and reliable.
Conduct risk assessments.	Consider the following types of assessments: -Gen AI Initial Use Case Assessment (See Appendix B). - AI Risk Threshold Analysis. - AI Risk Impact Assessment. -Algorithmic Impact Assessment	Human-based errors. Discrimination. Automation bias. Non-transparency. Lack of explainability. Lack of accountability. Threats to privacy.	Lawful, fair and responsible use. Transparency & accountability. Safe, secure, and resilient. Valid and reliable. Privacy and data protection. Accountability and responsibility.

8. EVALUATION OF RISKS/DETERMINATION WHETHER TO PROCEED			
Action or Control	**Action or Control**	**Action or Control**	**Action or Control**
Consider identification of risks to clients or customers if the Gen AI technology is deployed.	Consider risks to rights and freedoms to individuals. Is this a High-Risk Gen AI system with potential impacts to critical infrastructure? Health? Safety? Employment? Democratic or judicial process? Biometric identification or verification? Surveillance?	Discrimination. Lack of explainability. Threats to privacy.	Lawful, fair and responsible use. Fairness, inclusion, and non-discrimination. Explainable and interpretable. Public purpose and social benefit.
Can identified risks be accepted, reduced, mitigated or remediated by the agency in accordance with defined success criteria?	Determine a review period for reassessment.	Lack of accountability.	Lawful, fair and responsible use. Transparency & accountability. Safe, secure, and resilient.
	Determine key performance indicators or a cadence of review of success criteria for deployment of Gen AI technology in production.	Lack of accountability.	Lawful, fair and responsible use. Transparency & accountability. Safe, secure, and resilient. Valid and reliable.

8. EVALUATION OF RISKS/DETERMINATION WHETHER TO PROCEED			
Action or Control	**Action or Control**	**Action or Control**	**Action or Control**
	Determine tracking and monitoring mechanisms for managing risks and success criteria.	Lack of accountability.	Lawful, fair and responsible use. Transparency & accountability. Safe, secure, and resilient.
	Consider updating Incident Response Plan, Business Continuity Plan and Disaster Recovery Plan given certain safety and operational features of Gen AI technology. Consider whether back-up systems or datasets need to be generated/configured if the Gen AI technology fails or causes intolerable risk. Are necessary agency staff able to revert all control to human worker decision-making as needed?	Lack of explainability. Lack of accountability.	Lawful, fair and responsible use. Transparency & accountability. Security. Safe, secure, and resilient. Valid and reliable.
Can the integrity of decisions made or content produced by High-Risk Gen AI systems or integrity of outputs be confirmed or reviewed by state agency workers?	Does the Gen AI system flag or notify state agency workers of uncertainty in the answer/projected response to a series of data or information?	Human-based errors. Discrimination. Inaccuracies. Automation bias. Non-transparency. Lack of explainability.	Lawful, fair and responsible use. Purpose limitation. Transparency & accountability. Explainable and interpretable. Transparency & auditability.

8. EVALUATION OF RISKS/DETERMINATION WHETHER TO PROCEED			
Action or Control	**Action or Control**	**Action or Control**	**Action or Control**
	Does the Gen AI technology permit the state agency worker to override or manipulate the output or generated response?	Human-based errors. Discrimination. Inaccuracies. Automation bias. Non-transparency. Lack of explainability.	Lawful, fair and responsible use. Purpose limitation. Transparency & accountability. Safe, secure, and resilient. Accountability and responsibility.
Consider whether the testing environment accurately reflects the real-life production environment, including the client or customer datasets.	If unsure, consider piloting a proof-of-concept release of the Gen AI technology with production datasets to evaluate decisions-made or outputs by Gen AI technology for validity and success.	Human-based errors. Discrimination. Inaccuracies. Automation bias. Non-transparency. Lack of explainability.	Lawful, fair and responsible use. Purpose limitation. Transparency & accountability. Valid and reliable.

8. EVALUATION OF RISKS/DETERMINATION WHETHER TO PROCEED			
Action or Control	**Action or Control**	**Action or Control**	**Action or Control**
Are clients or customers able to opt-out of participation in the High-Risk Gen AI system and still obtain benefits/services offered/administered by the agency?	What alternatives are available to clients or customers if they opt-out?	Discrimination. Inaccuracies. Non-transparency. Lack of accountability. Threats to privacy. Threats to legitimacy and public trust.	Lawful, fair and responsible use. Purpose limitation. Transparency & accountability. Privacy and data protection. Fairness, inclusion, and non-discrimination.
	Consider notification of opt-out implications.	Threats to privacy. Threats to legitimacy and public trust.	Lawful, fair and responsible use. Purpose limitation. Transparency & accountability. Privacy and data protection. Fairness, inclusion, and non-discrimination.

9. REVIEW OF DECISIONS BY THOSE AFFECTED			
Action or Control	**Notes**	**Risk Addressed**	**Principles**
Has the agency identified the populations that will be involved with and/or impacted by the deployment of the High-Risk Gen AI system?		Discrimination. Lack of accountability. Threats to privacy. Threats to legitimacy and public trust.	Lawful, fair and responsible use. Privacy and data protection. Fairness, inclusion, and Non-discrimination. Public purpose and social benefit.
Consider assessing risk of those affected at population-level by conducting user-testing.		Discrimination. Lack of accountability. Threats to privacy. Threats to legitimacy and public trust.	Individual participation. Fairness, inclusion, and non-discrimination. Public purpose and social benefit.
Consider implications of known vulnerabilities of target populations.		Human-based errors. Discrimination. Inaccuracies. Threats to privacy. Threats to legitimacy and public trust.	Lawful, fair and responsible use. Fairness, inclusion, and Non-discrimination. Public purpose and social benefit.

9. REVIEW OF DECISIONS BY THOSE AFFECTED			
Action or Control	**Action or Control**	**Action or Control**	**Action or Control**
Consider formally documenting and publishing an FAQ about the High-Risk Gen AI system, including operational details of the decision-making model.	Include different components of the FAQ: - Dataset information. - Model information. - Human involvement in the High-Risk Gen AI system. - Inferences and projections of the High-Risk Gen AI system. - Impact of decisions-made by the High-Risk Gen AI system. - Ability to appeal decisions-made by the High-Risk Gen AI system.	Non-transparency. Lack of accountability. Threats to legitimacy and public trust.	Lawful, fair and responsible use. Transparency & accountability. Individual participation. Transparency and auditability.
Consider publishing outcomes of evaluation to procure or deployment of High-Risk Gen AI system.		Non-transparency. Lack of accountability. Threats to legitimacy and public trust.	Transparency & accountability. Individual participation. Accountability and responsibility. Transparency and auditability.

9. REVIEW OF DECISIONS BY THOSE AFFECTED			
Action or Control	**Action or Control**	**Action or Control**	**Action or Control**
Consider whether the agency should have a policy about publishing a report on agency development & procurement, deployment or use of High-Risk Gen AI system.	Consult with WaTech.	Non-transparency. Lack of accountability. Threats to legitimacy and public trust.	Transparency & accountability. Individual participation. Accountability and responsibility.
Did the agency address usability problems and test whether user interfaces served intended purposes?	Consider developing formalized feedback loop with affected populations. Consult target population to further understand potential or known vulnerabilities.	Discrimination. Lack of accountability. Threats to privacy. Threats to legitimacy and public trust.	Data minimization. Purpose limitation. Transparency & accountability. Individual participation. Public purpose and social benefit. Valid and reliable.

9. REVIEW OF DECISIONS BY THOSE AFFECTED			
Action or Control	**Action or Control**	**Action or Control**	**Action or Control**
Consider formal feedback channel on impact or for questions regarding High-Risk Gen AI systems from both internal and external participants.	Designate appropriate workforce personnel to manage the feedback channel.	Non-transparency. Lack of accountability. Threats to legitimacy and public trust.	Transparency & Accountability. Individual participation. Public purpose and social benefit. Fairness, inclusion, and Non-discrimination.
	Formalize how feedback will be ingested and considered by executive sponsors. Formalize how feedback will be utilized for ongoing review of High-Risk Gen AI system use.	Non-transparency. Lack of accountability. Threats to privacy. Threats to legitimacy and public trust.	Transparency & Accountability. Individual participation. Safe, secure, and resilient. Valid and reliable. Fairness, inclusion, and non-discrimination.
Consider redress process for individuals regarding decisions made about them by the High-Risk Gen AI system.	Clarify redress for updating/correcting client or customer data used in by High-Risk Gen AI System and/or the outcome/decisions made by the High-Risk Gen AI regarding the individual.	Non-transparency. Lack of accountability. Threats to legitimacy and public trust.	Transparency & accountability. Individual participation. Fairness, inclusion, and non-discrimination.

10. WEIGHING ADVANTAGES AGAINST KNOWN BIAS OR INACCURACIES			
Action or Control	**Notes**	**Risk Addressed**	**Principles**
Does the agency staff have insight into the model/dataset used for the Gen AI technology?	Did the agency account for unintended biases of the dataset through mitigation/remediation actions?	Discrimination. Automation bias. Non-transparency. Lack of explainability. Lack of accountability. Threats to privacy.	Lawful, fair and responsible use. Data minimization. Purpose limitation. Transparency & accountability. Explainable and interpretable. Transparency and auditability.
	Liaise with the agency data scientist and/or WaTech to evaluate potential bias of data model and/or dataset of Gen AI model.	Discrimination. Automation bias. Non-transparency. Lack of explainability. Lack of accountability. Threats to privacy.	Lawful, fair and responsible use. Data minimization. Purpose limitation. Transparency & accountability. Fairness, inclusion, and non-discrimination.
Consider perception of risk of identified Gen AI technology by target population or subpopulation.	Identify and document perceived risk. Document risk reduction strategies. Determine whether agency should accept, reduce, mitigate or remediate risks if development, r procurement, deployment, or use of Gen AI technology continues.	Discrimination. Inaccuracies. Automation bias Non-transparency. Lack of explainability. Lack of accountability. Threats to privacy.	Lawful, fair and responsible use. Data minimization. Purpose limitation. Transparency & accountability. Fairness, inclusion, and non-discrimination. Accountability and responsibility.

Appendix A

Sample Gen AI procurement questionnaire

1. What is the Gen AI Technology name, vendor and version?
    a. Gen AI Model Name, LLM Version (including number of parameters) and list ALL model names/owners for the solution or offering.
        i. Include the unique identifier or name assigned to the specific Gen AI model or service.
        ii. Identifiers should allow users to refer to and distinguish between different Gen AI models.
2. List Gen AI powered, or AI driven applications and/or product owners.
    a. Answer should include the name of the organization or entity responsible for creating or deploying the Gen AI model or service.
    b. Answer should help identify the source and accountability for the Gen AI system or technology.
3. Describe the product.
    a. Provide a concise summary of the Gen AI model's purpose, functionality, and key characteristics including model architecture, capabilities, and limitations.
    b. Provide a high-level understanding for users and government customers.
4. What are the Use Case(s) for the Gen AI technology?
    a. Describe the intended use or goal of the Gen AI model (e.g. image recognition, natural language processing, text summarization, data analysis)
5. What is the context, subject matter, or domain for which the Gen AI model is designed to operate effectively?
    a. Answer should help user determine if the Gen AI model is suitable for their specific use case.
    b. What data is the model trained on and what data will be input into the system when in use?
6. What decisions or outputs is the system utilized to make?
7. How is the input data gathered, how often is it updated, and are subjective inputs ever audited for consistency across data collectors?
8. Is the training data or decision or output algorithm available for examination by the agency and/or the public?
9. Has there been any public or community engagement used in selection or design of the system? If so, please describe this engagement.
10. Does law or regulation mandate any of the decision system criteria? If so, which criteria?
11. What quality control is in place to test and monitor for potential bias?
12. What performance metrics were selected to determine the models' effectives? What is the model optimizing for, and under what constraints?

13. Do the system's decisions intentionally differentially affect members of protected classes[1], such as selecting persons with disabilities for certain benefits?
14. Has the system been tested for unintended bias by the agency or an independent third party? If so, what were the results? Describe briefly the nature of the testing.
15. Has the system been tested for misuse potential, including for the production of non-consensual intimate imagery, child sexual abuse material, hateful or harmful content, offensive cyber capabilities, or other types of misuse? If so, what were the results? Describe briefly the nature of the testing.
16. Has the Gen AI technology produced known erroneous results and if so, what were those errors (including the results of any audits conducted to check for erroneous results)?
17. How are the outcomes of the Gen AI system explained to subject matter experts, users, impacted individuals, others?
18. In addition to any intentional differential effect on members of a protected class, are there other differential effects on protected classes as shown by comparison of the system's data to general census data or, where relevant, subpopulation data, such as the effect on justice system defendants of color as contrasted with all defendants? If audits have been performed to determine such differential effects, what were the results of those audits?
19. Can those affected by a High-Risk Gen AI system decision review and challenge the basis for that decision? If so, how, and what were the results of any such challenges?
20. Is the high-risk Gen AI system operated by a third party? If so, what rules govern such operation and what audits are conducted to ensure compliance?
21. What is the fiscal impact of the system, including initial cost, operating costs, and any cost savings established as flowing from use of the system?

[1]A protected class is a group of people who are protected from discrimination based on a shared characteristic. These characteristics include race, color, religion, sex (which includes pregnancy, sexual orientation, and gender identity), national origin, age, disability, and genetic information (including family medical history).

Appendix B

Sample Gen AI initial use case assessment

Below is a method to identify and assess whether your use case is appropriate for Gen AI technology and whether it is high-risk. Business needs or problems to be assessed should be explained along with context information such as a cost-benefit analysis, anticipated operational impacts, description of use case or decisions being made, the approximate size of impacted population, and the need or advantages of the Gen AI technology.

Issue	Question
Identify business need or problem	What is the business need or problem your agency is trying to solve?
Describe use case	What is the intended use or goal of the Gen AI model (e.g. image recognition, natural language processing, text summarization, data analysis)?
Cost-benefit analysis	What is the cost-benefit analysis of procuring, developing, or using a Gen AI system?
Operational benefits	What are the operational benefits of procuring, developing, or using a Gen AI system?
Effect on individuals	What is impact on individual rights and freedoms?
Risk of bias or errors	What is the data or training data involved in the Gen AI technology? Risk of harmful effects of bias content? Is there a perpetuation of stereotypes?
Complexity	What is Gen AI model generating or predicting?
	What are the anticipated outputs?
High-risk	Does the Gen AI system have potential impacts to critical infrastructure? Health? Safety? Employment? Democratic or judicial process? Biometric identification or verification? Perform surveillance?

To determine risk for further review and assessment, agencies should evaluate the following:

Effect on people

- Low: Gen AI technology does not impact legal rights or the provision of services or scrutiny that could lead to an impact on legal rights or services.

- Moderate: Gen AI technology impacts processing, relatively minor services or legal rights or financial impacts on individuals. Small number of impacted individuals.

- High: Gen AI technology can have a significant impact on the provision of services, financial impact, or legal rights. Large number of impacted clients.

Likelihood of bias or errors

- Low: Gen AI technology directly follows federal or state regulations or follows adopted policy or rule.
- Moderate: Gen AI technology developed with disclosure of information used and the model has been tested for bias and inaccuracy.
- High: Gen AI technology developed without disclosure of information used or the model inputs, training data or outputs created and has not been tested for bias or accuracy.

Complexity

- Low: Simple model or decision rule.
- Moderate: Simple calculation of existing data elements (i.e., a weighted average).
- High: Complex predictive modeling, deep learning, etc.

Matrixes could be used to determine a rating such as:

	Low impact	Moderate Impact	High Impact
Low likelihood	5	4	4
Moderate likelihood	4	3	3
High likelihood	4	2	1

Risk and complexity rating can be used to determine the type of review, for example:

- Low: Generative AI questionnaire.
- Moderate: AI risk analysis or formal outcome analysis.
- High: Gen AI impact assessment by professional or third party.

Generative AI Technology Evolution

WaTech acknowledges that the field of generative artificial intelligence is rapidly evolving.

The guidance provided serves as a tool for agencies, drawing from the recommendations outlined in the White House's Blueprint for an AI Bill of Rights and the National Institute for Science and Technology's (NIST) AI Risk Management Framework and other work the state has undertaken in regard to automated decision systems. It is rooted in a foundation of guiding principles, providing a framework for agencies to navigate the complex landscape of generative AI technology. It is important to recognize that this guidance will evolve over time, reflecting the advancements and changing landscape of technology and governance. While it is to be used as a best practice, it should not serve as a substitute for regulations or agency policies that may be in place. Instead, it should be adopted in alignment with other relevant resources and tools, ensuring a comprehensive and well-rounded approach to addressing the challenges and opportunities presented by generative artificial intelligence.

Contact

Questions regarding this document can be directed to:

Washington Technology Solutions

ai@watech.wa.gov

References

Executive Order 24-01 on Artificial Intelligence Section 3. requiring this deliverable reads as follows:

By September 2024, WaTech, in collaboration with the Department of Enterprise Services (DES), shall issue initial guidelines for public sector procurement, uses, and ongoing monitoring of the use of generative AI technology. The guidelines should build on guidance from the White House's Blueprint for an AI Bill of Rights and the National Institute for Science and Technology's AI Risk Management Framework and address safety and effectiveness, algorithmic discrimination, data privacy, and cybersecurity. These guidelines shall include a requirement that any vendor providing a High-Risk Generative AI System to an agency certify that the vendor has implemented an AI governance program consistent with the National Institute of Standards and Technology Artificial Intelligence Risk Management Framework. WaTech and DES shall make the guidelines publicly available via posting on relevant agencies' websites.

References and materials used to develop this document include the following:

- White House AI Bill of Rights Intro: https://www.whitehouse.gov/ostp/ai-bill-of-rights/
- White House AI Bill of Rights: https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf
- NIST AI RMF: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf
- NIST AI RMF: Gen AI: https://airc.nist.gov/docs/NIST.AI.600-1.GenAI-Profile.ipd.pdf
- CA Gen AI Guidelines for Public Sector: https://cdt.ca.gov/wp-content/uploads/2024/03/3a-GenAI-Guidelines.pdf
- CA Gen AI Fact Sheet: https://www.documents.dgs.ca.gov/dgs/fmc/pdf/std1000.pdf
- Canada ADS Directive: https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592
- World Economic Forum – AI Government Procurement Guidelines: https://www3.weforum.org/docs/WEF_AI_Procurement_in_a_Box_AI_Government_Procurement_Guidelines_2020.pdf
- World Economic Forum - AI Procurement in a Box Toolkit: https://www3.weforum.org/docs/WEF_AI_Procurement_in_a_Box_Workbook_2020.pdf
- AI Institute Algorithmic Impact Assessment Report: https://ainowinstitute.org/publication/algorithmic-impact-assessments-report-2