

Guidelines for Deployment of Generative AI

December 2024

WaTech
Washington Technology Solutions

Table of Contents

Executive summary.....	3
Background.....	5
Goals of using Generative AI technology.....	6
Risks.....	6
Guidelines for analysis of impact.....	8
Identifying vulnerable communities.....	8
Engage interested parties.....	9
Prior to engagement.....	11
During engagement.....	12
Post engagement.....	13
Community engagement checklist.....	13
Key principles.....	13
Engagement of Tribal Governments and Communities.....	17
Criteria for evaluating equitable outcomes.....	19
Accountability.....	19
Evaluate potential biases.....	20
Privacy, security and data considerations.....	22
Privacy.....	22
Security and data.....	23
Equity of access.....	24
Inclusivity in design.....	25
Quality of outcomes.....	26
Mitigation of harm.....	26
Long-term sustainability.....	27
Evaluation and monitoring.....	28
Conclusion.....	29
References.....	30

Executive summary

This guidance was created by Washington Technology Solutions (WaTech) in collaboration with the Washington State Office of Equity (OOE) and other partners pursuant to the [Governor's Executive Order 24-01](#). The purpose is to provide guidance that state agencies can use to analyze the impact of generative artificial intelligence (GenAI) technology on vulnerable communities.

Per Executive Order 24-01 by December 2024, WaTech, in consultation with the Office of Equity, community members, Tribal governments, subject matter experts, and other relevant state agencies, shall develop guidelines for agencies to analyze the impact that adopting generative AI technology may have on vulnerable communities, including criteria to evaluate equitable outcomes in deployment and implementation of high-risk use cases. These guidelines and criteria shall inform whether and how an agency deploys a particular AI System. WaTech shall periodically revise the guidelines and criteria, when appropriate, and establish a consultative process with academia, industry experts, and representatives from advocacy organizations that represent communities that are disproportionately vulnerable to being harmed by algorithmic bias. WaTech and the Office of Equity shall make the guidelines publicly available by posting on their respective agency's website.

Understanding the impacts of adoption and deployment of GenAI on vulnerable populations is important because it allows for effective and equitable solutions that provide equitable outcomes. This guidance is an important part of Executive Order 24-01, which emphasizes safe, ethical, and effective deployment of GenAI. Sound strategies that account for the experiences and needs of those historically excluded from new strategies and solutions have a meaningful impact on reversing harms to vulnerable populations and ensuring equitable outcomes. Identifying vulnerable populations and risk management capacity is crucial for understanding and mitigating risks and harms while increasing public trust.

Community engagement is a key part of a human-centered strategy for addressing disproportionate impacts of adopting new technologies. Analyzing data and measuring outcomes about the impacts of policies and practices that accompany innovative solutions on vulnerable groups is essential to understanding the realities these communities face. The World Health Organization defines vulnerability as *the conditions determined by physical, social, economic, and environmental factors or processes which increase the susceptibility of an individual, a community, assets, or systems to the impacts of hazards*.

GenAI has the potential to improve efficiency and reduce costs while enhancing equity and delivering more effective public services. While this technology offers potential for innovation, it must be deployed with care to avoid reinforcing biases or creating unintended harm.

This report highlights:

Gen AI risks

The use of GenAI presents potential challenges such as:

- Inaccuracies and biases.
- Lack of transparency and accountability.
- Privacy and cybersecurity concerns.
- Threats to public trust due to non-human decision-making.

Agency guidelines

To reduce the potential impact on vulnerable communities, agencies must first understand both positive and negative impacts of GenAI deployment. Agencies should:

1. Identify vulnerable communities affected by GenAI: These include groups such as the elderly, individuals with disabilities, ethnic and racial minorities, and those with limited digital literacy.
2. Engage communities through culturally appropriate and inclusive methods: Agencies should prioritize early and transparent communication, tailoring outreach efforts to address the cultural and linguistic needs of diverse communities. By building partnerships with community representatives and Tribal governments, they can ensure inclusive decision-making and foster trust. Agencies must measure the outcomes of their engagement initiatives and adapt their processes based on community feedback.
3. Address and mitigate risks while fostering equity: The state must address the importance of minimizing data collection, ensuring only essential information is gathered to respect individual privacy and limit exposure to potential risks. Transparency in data usage is equally critical, requiring clear and open communication about how data will be used and why it is being collected. To uphold these principles, maintaining robust security practices is paramount, safeguarding data against misuse and fostering trust in the systems and processes that handle sensitive information.
4. Develop monitoring systems for continuous evaluation: Agencies must prioritize monitoring to identify and mitigate potential biases in data and AI models, ensuring equitable outcomes for all groups, especially vulnerable populations. Biases can emerge at any stage, from data collection and sampling to model training and interpretation. To address this, agencies should implement ongoing monitoring processes, such as exploratory data analysis, regular audits of AI models, and cross-validation to detect and correct biases.

Equity of access and outcomes

Barriers such as digital illiteracy and resource inequities must be proactively addressed. Agencies should implement human-centered design and adopt mechanisms for continuous community feedback.

Long-term sustainability

Agencies should incorporate sustainability principles into their strategies, ensuring that benefits of GenAI technologies are equitable and enduring over time.

Evaluation and monitoring

It's important to involve people in the evaluation process. Agencies need to provide enough resources to regularly check, assess risks, and improve GenAI systems. By fostering collaboration, transparency, and proactive risk management, agencies can build trust and achieve positive outcomes for all communities.

Background

The Executive Order (EO) recognizes that GenAI is a transformative technology which has the capacity to improve the way 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 (GenAI) 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 have; along with the Washington State Department of Health's (DOH) and King County Community Engagement guides, been 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."

Goals of using Generative AI technology

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

- **Efficiency:** GenAI may assist governmental entities in streamlining processes and making administrative decisions more quickly and efficiently.
- **Economical stewardship:** GenAI 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 GenAI may introduce and reinforce bias, it also potentially provides an opportunity to address biases.
- **Effectiveness:** Improving delivery of public services: GenAI may help the public receive improved and more accessible services.

Risks

There are several known risks of using GenAI technology listed below. The way that AI systems are developed and implemented poses challenges for how risks can be identified and managed. The *Implementing Risk Assessments for High-Risk AI Systems* provides additional information regarding risks that should be understood before deployment of any GenAI tools. Understanding these challenges, the specific types of risk posed by AI systems, and leveraging existing processes can all help agencies innovate responsibly and effectively manage risk.

- **Inaccuracies:** GenAI 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 GenAI 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 GenAI is being used, how that system operates, and the impacts of the system on individuals and society.
- **Lack of explainability:** GenAI 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 GenAI tool is procured through a third-party vendor. Even the developers of GenAI 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 GenAI outputs may not have the ability to meaningfully challenge a system's decisions. Governmental entities that procure or use GenAI 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 GenAI 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 GenAI 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:** GenAI 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 improperly reviewed resources or systems.

Guidelines for analysis of impact

To lessen the impact to vulnerable communities, agencies must first understand both positive and negative impacts of deployment. Without understanding and addressing the impact, the potential remains to increase disparities experienced in the determinants of equity below.



Equity in State & Local Practices



Early Childhood Development



Quality Education



Equity in Jobs & Job Training



Healthy & Human Services



Food Systems



Parks, Recreation & Natural Resources



Healthy Built & Natural Environments



Transportation & Mobility



Economic Justice



Strong, Vibrant Neighborhoods



Housing & Home Ownership



Community & Public Safety



Equity In Justice Systems & Laws



Digital Equity

Determinants of equity from King County and Washington State Office of Equity

Identifying vulnerable communities

To identify vulnerable communities, consider that there are established examples of groups and communities that are particularly vulnerable and disproportionately affected by the impacts of technology. These communities are population groups that often experience harm, discrimination, or disadvantage due to various factors such as social, economic, geographic location, or physical circumstances. Due to those realities these groups may experience:

- Increased health problems and lack of access to quality care.
- Socioeconomic harms due to inequities in finances, education, geographic location, housing, and underemployment.

- Distrust in accessing government services and programs.
- Limited English proficiency in order to give informed consent or otherwise communicate effectively.
- Limited digital literacy.

Included but not limited to within these communities are elderly individuals, children, individuals with disabilities, members of ethnic and racial minority populations, LGBTQIA+ individuals, low-income populations, those who are unhoused, refugees, immigrants or migrants, and those experiencing some form of abuse.

Equity in deployment must consider the status of individuals who belong to underserved communities that have been denied fair treatment, such as Black, Latino, Indigenous, Asian Americans and Pacific Islanders, other persons of color and those that live in underserved or unserved areas of the state. Vulnerability is determined by historical, political, cultural, institutional, and natural resource processes that shape people's lives and may produce a range of underlying drivers of vulnerability.

State agencies should use this guidance to:

- Identify vulnerable populations.
- Engage interested parties in communities and Tribal governments.
- Evaluate potential biases in deployment through engagement and data analysis.
- Mitigate risks to agency and communities impacted by use of GenAI.
- Perform ongoing monitoring and evaluation of identified impacts.

Engage interested parties

It is important to engage with community representatives to understand specific needs and concerns and experiences. Community engagement is fundamental to protecting the interests of residents engaging in Washington state services. Community engagement efforts can advance digital equity, promote social connection, strengthen partnerships, and build relationships with the communities served by Washington state agencies. This guidance is intended to help state agencies analyze the impacts of GenAI adoption on vulnerable communities by providing a consistent approach that:

- Leads with equity as a core principle.
- Ensures culturally and linguistically appropriate engagements.
- Promotes partnership with community.
- Uses feedback to improve the effectiveness of state services.
- Promotes partnership with impacted communities from the beginning.

- Avoids using an informing model of engagement whereas information that has been decided is presented to impacted communities with little or no consideration of that community's input.
- Demonstrates cultural humility by respecting differences in perspectives and experiences.
- Considers the needs of all residents.

The Office of Equity has developed the **SPICE** framework to center community experience in programs and services as well as create a path towards intentional and comprehensive community interaction. Review the model below and reflect on how your agency can engage with community on a deeper level beyond service. How can you embody these five principles?

S - Seek new partners: Learn and understand how to interact with other communities in a culturally appropriate and respectful manner. Seek out community instead of waiting for the public to reach out to government.

P - Partner: Work in a shared decision-making model. Being a partner means sharing in investment, benefit, and risk even if they don't look the same on either side.

I - Inform: Provide updates in an accessible manner and share information that will reach different communities.

C - Connect, collaborate, compensate: Make contact to build or strengthen relationships. Focus on co-creation and compensate people in the community for their ideas, time, and energy instead of extracting insights.

E - Engage: Interact with others through activities or events and listen to understand the perspectives of community members.

For more information on community engagement best practices, contact the [Office of Equity](#).

Successful engagements can be accomplished in several ways including:

- **Targeted engagement based on impact and identity:** This considers groups and communities that have been or may be adversely impacted by the deployment of GenAI tools and the specific impacts to their respective community.
 - Utilize indicators for analyzing and defining vulnerable communities such as age, income, education, ability, ethnicity and social caste.
 - Other indicators may include demographics such as housing security, mobility, availability of health services, environmental hazards, business/job opportunity, available public and private utilities, and social services.
- **Geographical engagement:** Identify and prioritize regions of the state or Tribal land boundaries where the impact of GenAI tools to groups and communities should be

prioritized, based on specific criteria. Such criteria may include mountainous regions, or remote and rural locations.

- **Community participatory engagement:** In collaboration with impacted communities, develop an action plan to engage with community early on. This plan should include clear goals as decided with the community, defined tasks, timelines, and responsibilities for each part of the plan.
 - Involve community members and organizations, Tribes, advocacy groups, relevant state agencies and industry partners in discussions about potential impacts to vulnerable communities.
 - Develop a process to evaluate and measure engagements and results.
 - Assess the effectiveness of the engagements and create feedback loops to ensure ongoing dialogue and adjustment based on community input.

Prior to engagement

- **Review and implement key principles of community engagement**
 - Learn about the community.
 - Understand any biases, assumptions or stereotypes that may be present.
 - Understand where individuals or organizations have historically held benefits and privileges that negatively impact the community.
- **Identify the community and contact early in the development process**
 - Leaders.
 - Members.
 - Partners.
 - Other interested parties.
- **Choose and engagement method**
 - Ensure engagements are done in a place, date and time that are accessible and culturally and linguistically appropriate.
- **Plan on how to define and measure success**

In every software application that incorporates AI, you should identify performance measures to track over a defined period (i.e. 12-months following deployment). This will help the agency review the benefits of the GenAI system and compare these to the harms introduced or perpetuated by the use of the system. For example:

- When procuring a translation management program that leverages GenAI because the agency produces enough content that translation contractors can't keep up, then the following should be considered:

- Starting from proposal through procurement and contract close estimate the average time to complete a standard translation using translation contractors through the Department of Enterprise Services (DES) as a baseline.
- Track an average time to complete translations using the translation management program and GenAI.
- After the 12-month tracking period, compare the two values and determine if the GenAI system has created the desired solution.

Impact is often harder to measure if not considered at the beginning of efforts. Make sure to discuss not only what the tool does and how to measure performance, but also how it impacts your customers, workforce, and community.

Using the same example above, the agency can then measure how many more languages materials can be provided because turnaround time is greatly reduced.

Providing translations in more languages helps in engaging with communities that are often left behind due to language barriers.

The agency might consider measuring how much staff time can be re-allocated to other projects. Another option is to monitor how the agency is supporting employees with upskilling or reskilling in response to the positive or negative impacts of automation on certain job roles.

During engagement

- **Continuously assess engagement on methods and approach**
 - Check-in regularly with participants to ensure engagement is being well received and culturally considerate.
 - Based on evaluations of the engagement or during the project make needed adjustments to structure, membership and process.
- **Identify previously unknown barriers**
 - As new barriers are learned, acknowledge said barrier and plan to remove or lessen the barrier so as to ensure equitable participation.
- **Listen to understand**
 - All voices are given equal value in the engagement.
 - Limit assumptions about people, experiences and access.
- **Honesty and transparency**
 - Be honest about what can and cannot be done.

- Communicate why information is being asked for and how that information will be used.

Post engagement

- **Gather feedback**
 - Seek feedback from community and other partners involved and use those perspectives to evaluate overall success and continuously improve.
- **Share results**
 - Share outcomes and next steps with community and other partners in an accessible manner.
- **Evaluate ways to improve and stay connected**
 - Improve long-term relationships by evaluating and improving processes.
 - Stay connected with community partners to continue to increase trust.
 - Be clear about how the information given is being utilized throughout the project.

Community engagement checklist

Engagement with communities should be done in a thoughtful, responsible and culturally respectful manner. The continuum below provides methods on how such engagements can occur in a manner that promotes community-driven engagements. Every effort should be made to consider cultural contexts when engaging with various communities to ensure that GenAI applications respect cultural contexts and do not perpetuate stereotypes; and consider community values and norms in the design and implementation process.

Key principles

To build strong relationships that can help analyze impacts to communities it is important to recognize the diversity of those communities. Diversity must be understood as more than what is visually seen and consider the intersectionality of each members lived experiences and cultural customs and norms. Taking a humble approach with all groups allows for learning about the nuances present that make the community who they are. Following key principles of engagement helps to build trust with community members and partners.

Key engagement principles

<p>Understand the community</p>	<p>It's important to understand the community's culture, norms, values, power and political structures, economic conditions, social networks, demographic trends, and history.</p> <p>Be aware of the community's history of collaboration with other agencies and partners, including the barriers they have faced.</p>
<p>Understand yourself and organization</p>	<p>Before engaging with a community or potential partner, check your own biases, privileges, and limitations:</p> <ul style="list-style-type: none"> • What assumptions or stereotypes do you hold about this community? • How could those assumptions negatively impact your interactions or efforts? • Do your cultural norms and values align with those of the community? <ul style="list-style-type: none"> ○ If not, how will that be accounted for? • Are you the most effective person to be leading this engagement effort, or should you work with or through partners?
<p>Allow community to self-identify</p>	<p>Cultural identities are dynamic, how people identify is a personal choice. People are both individuals and members of various groups.</p> <p>Some communities and individuals may have intersectional identities, meaning they are members of more than one group that has been historically marginalized or oppressed.</p>
<p>Emphasize historically unheard perspectives</p>	<p>Listen with the intent to act on the needs expressed by the community.</p> <p>Recognize that individuals have different perspectives, and some have been historically underrepresented.</p>
<p>Compensate for the sharing of lived experiences</p>	<p>Implementation and deployment of policies and projects should be informed by those that have an interest in the outcomes.</p> <p><u>Senate Bill 5793</u>, the legislation around Community Compensation, is meant to direct agencies to consult</p>

	<p>and compensate community members who have been historically excluded from the process.</p> <p>The value community brings to projects can be shown through compensation, reimbursement, or support for their priorities.</p>
Avoid tokenism	<p>Individuals should never be expected to speak on behalf of, or represent, an entire community.</p> <p>Communities themselves are not monolithic. While there are shared experiences and commonalities, each individual within a community has a unique story, identity, beliefs, and values.</p>
Acknowledge community strengths and culture	<p>Every community has strengths that should be acknowledged.</p> <p>Communities that experience the greatest inequities should never be undervalued.</p>
Proactive not reactive	<p>Reach out to and collaborate with potential partners and community members as early as possible.</p> <p>Doing so creates an environment of co-creation through input of design rather than feedback of solutions that have been decided without the input of community.</p>
Consistent and clear communication	<p>Continually use two-way communication methods that partners or community members are familiar with using and are accessible.</p> <p>Use plain language in communications.</p>
Transparency	<p>Be honest about purpose and how the input received will be utilized.</p> <p>Follow through on commitments.</p>
Meet them where they are	<p>Be flexible in your approach. Go to the community—where members gather—and work to build relationships and trust.</p> <p>Provide accessible non-traditional ways to connect with communities.</p>

Co-creation with the community is a valuable strategy to ensure that the harmful impacts of adoption of new technologies are minimized, and gaps are readily identified before launch.

Prioritizing co-creation with the community helps agencies achieve relevant, effective, and sustainable outcomes that benefit the people they aim to serve.

Other benefits of a co-creation model include:

- Promoting innovation by considering diverse perspectives and deploying collaborative problem solving.
- Empowering communities to be a part of the solutions they seek including providing opportunities for skill development and capacity building within the community thus enhancing their ability to address future challenges.
- Adaptability as the solutions can be improved over time based on ongoing involvement and feedback of the community.
- Improved relationships between state government agencies and the community.

Historically, engagements between government and the public mostly informs residents, thus limiting community input from different groups or other interested parties. Agencies should understand the dynamics that are present in various types of engagement. The Community engagement nexus below can help understand where power can be shared in co-creation depending on the engagement model utilized.

Community engagement nexus					
	Communicate	Consult	Involve	Collaborate	Empower
Lead by	Led by state	Led by state	Led by state	Co-led by state and community	Led by community
Purpose	<ul style="list-style-type: none"> • Provide information. • Address immediate need. 	<ul style="list-style-type: none"> • Receive and respond to feedback. • Share progress of a program. 	<ul style="list-style-type: none"> • Ensure the needs of community are considered. • Understand the complexity of problems and provide pertinent solutions. 	<ul style="list-style-type: none"> • Partner with community in developing solutions that work for them. • Share decision making power. 	<ul style="list-style-type: none"> • Support community in developing solutions that impact their lives and respect cultural norms and practices.
Comm. Style	One-way	One-way	Two-way	Two-way	Two-way

Process	<ul style="list-style-type: none"> •Community meetings •Social media •Internet •Town halls 	<ul style="list-style-type: none"> •Focus groups •Surveys •Interviews with interested parties 	<ul style="list-style-type: none"> •User testing •Governance and steering committees •Advisory groups or boards 	<ul style="list-style-type: none"> •Coalition building •Government to Government partnerships •Public and private investments and partnerships. 	<ul style="list-style-type: none"> •Community immersion and mobilization.
Role	To keep interested parties informed.	Listen to community feedback.	Incorporate feedback into project or decision.	Work together to incorporate input from the beginning rather than feedback.	Implement community led projects or decisions.
When to Use	Information provided regularly dependent on urgency, regulatory or legal responsibility.	When improving existing service or program.	Needing community perspective and buy-in.	Whenever community has a strong interest in outcomes and can participate in a non-exploitive manner.	Community owned projects as part of a long-term relationship strategy.

Engagement of tribal governments and communities

Washington state is home to 29 federally recognized Indigenous tribes, each with their own Tribal governments. Tribal governmental and business enterprises are creating tens of thousands of jobs and investing billions of dollars in the state’s economy.

Each of these Tribes exercise tribal sovereignty. Sovereignty is a legal term meaning the right of a people to self-govern within jurisdictional borders. Tribal sovereignty recognizes Indigenous tribes as having the political status of nations. Sovereignty recognizes that these nations are geographically located within the territorial boundaries of the United States.

As sovereigns, Tribal nations have a government-to-government relationship with the two other sovereign governing bodies in the U.S. – the federal and state governments. Tribal law, federal law and state laws define government responsibilities, powers, limitations and

obligations. Tribal sovereignty allows tribal nations autonomy to govern, exercise jurisdiction and protect and enhance the health, safety and welfare of tribal citizens within tribal territory.

Indigenous knowledge and perspectives are essential for addressing challenges like social inequity. The deep understanding of ecosystems, sustainable practices, and community-based solutions offer valuable insights for navigating a rapidly changing world.

When looking at the potential impacts on Tribal governments and communities, agencies should be aware of:

- Creation of, or updates to policies, programs, and projects that will impact Tribes or Tribal communities may require a formal consultation process.
- Agencies should work with their Tribal relations staff to ensure engagement with Tribal governments and communities is done so in a manner that respects the sovereign status of the tribe.
- [RCW 43.376 - Government-to-Government Relationship with Indian Tribes \(Full RCW as PDF\)](#)
- [Centennial Accord](#)
- [Out of State Accord](#)
- [Millennium Agreement](#)

Other considerations agencies should be aware of when engaging with Tribal governments and communities include:

- Respecting and honoring the unique nature of Indigenous cultures which are diverse in nature and reflect the history and traditions of each community.
- Understanding the importance and historical relevance of storytelling.
- Honoring of spiritual beliefs held by each unique Tribe and its members.
- The opportunity to learn from Indigenous languages and traditions.

Criteria for evaluating equitable outcomes

Accountability

To provide services that are a net benefit for all, accountability in deployment of any technologies meant to enhance the digital experience of the user is a critical component of reducing harms. The accompanying *Framework for Accountability in Generative Artificial Intelligence for Washington State Agencies* outlines five pillars for accountability which are:

1. **Governance:** Creating effective operational structures around the use of GenAI.
 - Clear goals.
 - Roles, responsibilities, and monitoring.
 - Values.
 - Community involvement.
 - Transparency.
2. **Equity:** Centering the consideration and protection of vulnerable and underserved communities.
 - Emphasize human involvement and human impacts.
 - Accessibility and plain language.
 - Pro-Equity and Anti-Racism.
 - Generative AI images or video.
 - Generative AI usage transparency.
 - Data source accountability.
 - Bias detection and mitigation.
 - Trust and reputation of government.
3. **Data:** Outlining the responsibility for data sourcing and privacy protections.
 - Identify data sources.
 - Identify data pathways.
 - Hallucination tracking.
 - Responsible.
4. **Performance and Outputs:** Understanding risks and the importance of assessing the performance and impacts of outputs.
 - Measure bias.
 - High risk.

- Evaluate the quality of inputs.
- 5. **Continuous Monitoring:** Structuring monitoring plans to ensure continuous equitable and responsible use of GenAI.
 - Planning.
 - Traceability.
 - Opt-out.
 - Public feedback.
 - Transparency structure.

This enterprise-level framework should inform the agency-level policies.

Evaluate potential biases

Agencies should analyze data sources for biases that may affect the outcomes for vulnerable groups. Bias in data is a statistical distortion that can occur at any stage of the analysis. Systemic bias can be present in AI datasets and accepted organizational practices, and processes.

Examples of how this shows up includes how data is measured, sampled and observed in the process. Training GenAI systems with the wrong information can lead to biases showing up in data analysis such as those shown in the table below.

"Automated systems are not inherently neutral. They reflect the priorities, preferences, and prejudices - the coded gaze - of those who have the power to mold artificial intelligence.

We risk losing the gains made with the civil rights movement and women's movement under the false assumption of machine neutrality. We must demand increased transparency and accountability." - Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. Summary based on research by Joy Buolamwini

Confirmation Bias- bias that is supported by data that supports a pre-determined idea.

Availability Bias- reliance on data that is readily available but may not be robust in terms of how it was collected.

Selection Bias- targeted collection of data depending on factors such as when the data was collected. Often used to predict outcomes without accounting for changes in time.

Optimist/Pessimist Bias- analysis of data to support hopeful or negative outcome that supports other objectives.

Terminology Bias- GenAI models can introduce bias in analytics when models are trained on public data that uses terminology that differs from one organization to another.

POTENTIAL HARMS FROM ALGORITHMIC DECISION-MAKING





INDIVIDUAL HARMS			COLLECTIVE SOCIAL HARMS
ILLEGAL DISCRIMINATION	UNFAIR PRACTICES		
HIRING			LOSS OF OPPORTUNITY
EMPLOYMENT			
INSURANCE & SOCIAL BENEFITS			
HOUSING			
EDUCATION			
CREDIT			ECONOMIC LOSS
DIFFERENTIAL PRICES OF GOODS			
LOSS OF LIBERTY			SOCIAL STIGMATIZATION
INCREASED SURVEILLANCE			
STEREOTYPE REINFORCEMENT			
DIGNATORY HARMS			

Chart Contents Courtesy of Megan Smith, Former CTO of the United States

Agencies should be aware that many of the biases currently exhibited by GenAI produce output, content, and decisions that are harmful to groups that are already marginalized due to protected characteristics, including race, gender, sexual orientation, and disability, among others.

Some examples of how these harms show up include:

- Large language models exhibiting racial and gender biases that systematically disadvantage racial minorities and women.
- When summarizing and ranking resumes, AI tools used for automated screening can exhibit bias against resumes that imply the candidate has a disability.
- Generated images causing perpetuation of stereotypes and skewed perceptions of marginalized people and groups.
- Image generators lacking accurate representations of diverse relationships therefore perpetuating Euro-centric beauty standards.
- Data for policing models leading to a focus on specific historically marginalized neighborhoods,
 - This focus increases police presence and a potential for more arrests.
 - New data then used to further perpetuate focus on those neighborhoods.
- Recommender systems may prioritize popular views and not display minority views.

Ways to mitigate these biases include:

- Identifying the source of bias and any ways to counter those biases. Confirm the source is reputable.
- Conduct exploratory data analysis to detect potential biases.
- Cross-validation of GenAI models. User testing.
- Creating high quality data sets that are open to the public.
- Using different windows of time for creating training models that account for cyclical patterns.
- Conducting audits of GenAI models to ensure they do not reinforce existing inequalities.
- Accessibility and usability
 - Assess whether the technology is accessible to all segments of the community, including those with disabilities.
 - Evaluate the usability of the technology for individuals with varying levels of digital literacy.

Privacy, security and data considerations

Privacy

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 GenAI.

Security and data

Deploying GenAI comes with several security considerations to ensure the technology is used safely and responsibly. The [Office of Cybersecurity](#) and [Office of Privacy and Data Protection](#) provide information for agencies on how to best detect and prevent threats; and how to use and protect data responsibly.

As agencies look to AI tools for efficiencies in service delivery it should be noted that there are key security factors to consider such as:

Key security factors	
Factor	Action
Data Privacy and Protection	<ul style="list-style-type: none"> • Ensure that sensitive data used by GenAI models is encrypted and securely stored. • Implement access controls. • Stay informed about and comply with data protection policies and guidance.
Threat detection and response	<ul style="list-style-type: none"> • Use GenAI to enhance threat detection capabilities by analyzing patterns and anomalies in data.

	<ul style="list-style-type: none"> • Develop automated response systems to quickly address any detected threats. • Conduct regular audits to ensure compliance with security protocols.
<p>Training for users</p>	<ul style="list-style-type: none"> • Educate users about the potential risks and best practices for using GenAI. • Promote awareness of phishing attacks and other social engineering tactics that could exploit AI systems. • Train users on how to identify and mitigate biases in GenAI models. • Discuss various use cases and best practices for deploying GenAI in different scenarios.
<p>Security of AI models</p>	<ul style="list-style-type: none"> • Protect AI models from tampering and unauthorized access. <ul style="list-style-type: none"> ○ Agencies that do not have direct access to AI data models should consult with the AI developer to ensure protection from tampering and unauthorized access. • Use techniques like model encryption and secure areas to safeguard the integrity of the models. • Implement mechanisms to make GenAI decisions transparent and understandable to users. • Provide clear explanations for how the AI reaches its conclusions to build trust and accountability.

Equity of access

When individuals or groups of people are excluded or lack the knowledge, income, equipment, or training necessary to participate fully in public affairs, barriers to access must be overcome in order to ensure fairness.

Equity of access is the action of intentionally identifying opportunities to bridge the gap between those who have access to digital technologies and those that don't. This is achieved through human-centered, community-based participative design from the start, and adoption of mechanisms that seek regular feedback from the community to ensure that equitable access is maintained.

Rather than waiting for poor results or broken technology to drive equitable access, agencies are proactive in identifying current barriers and developing actionable and time-bound plans to address them.

Examples of questions to consider:

- Are vulnerable communities able to access the GenAI technology in an equitable manner?
- What measures are in place to bridge the digital divide?
- Is this technology an opportunity to bring technology and systems to communities that don't currently have access?
- What is preventing interested parties from being able to leverage technology?
- Have you worked with your agency IT Accessibility Coordinator to ensure that the agency accessibility plan includes support for complaint resolution?

Potential actions to take:

- Identify and document technology barriers and their impacts.
- Incorporate clear and measurable performance standards that include the outcomes to remediation of identified barriers and challenges.
- Prioritize what is learned about inequities of access and allocate resources as needed to eliminate them and achieve equity of access.
- Commit to ensuring state services are accessible regardless of access.
- Reducing gaps in access to state services where possible.

Inclusivity in design

Inclusive design aims to bring communities into the process of designing technology at the beginning, not after they've already been built. Being able to do this requires that state agencies become more aware of and unlearn harmful practices and behaviors that are still socially acceptable to create psychologically safe environments for people to speak up and be truthful in their responses.

Examples of questions to consider:

- Were community members involved in the design process from the early stages?

- Are the perspectives of diverse interested parties considered in development?
- Are there mechanisms for interested parties to provide meaningful feedback throughout the design process?
- Are the algorithms used in generative AI transparent?
- Is there a clear process for accountability if the technology causes harm?

Potential actions to take:

- Develop a community engagement model that offers an inclusive array of education and outreach opportunities that focus on moving forward while understanding how history has played a role in the current state.
- Invest in human-centered design practices and leverage existing policies so that agencies understand their purpose and are using them when developing new technologies.
- Develop a process to collect information about harmful impacts and ways that the agency will mitigate them.
- Share digital resources and knowledge and seek input from interested parties outside of the government ecosystem, to develop a collective approach to solving digital equity issues.

Quality of outcomes

The outcomes of efforts should not create more barriers and harms to the communities served by state agencies. To decrease impacts to vulnerable communities' agencies should consider the following questions:

- Do the outcomes produced by GenAI positively affect vulnerable communities?
- Does the agency understand where the disparities exist and how pervasive they are?
- Have metrics been identified to assess the quality and relevance of the AI-generated outputs?
- Do outcomes address "isms" such as racism, sexism, ableism as well as bias that may increase access issues and inequities?

Mitigation of harm

Mitigating harm recognizes that while it is not feasible to expect no harm to come from Washington state services and the data collected by those services, there must be controls and mechanisms put into place that reduce and mitigate harm where possible. When assessing new technologies, especially those that rely on large data sets like Large Language Models, it is imperative that the assessor take a risk-based approach in their assessment.

Examples of questions to ask:

- What strategies are in place to mitigate potential negative impacts on vulnerable communities?
- Are there support systems to address adverse effects if they occur?

Potential actions to take:

- Incorporate risk and accountability assessments into technology procurement processes.
- Offer a tool kit that provides guidance on how to spot and mitigate potential harms.
- Develop a long-term roadmap and set of actionable items that aim at addressing root causes of bias within the system itself.
- Balance state protections, sovereignty, and autonomy for marginalized individuals with the goal of gradually dismantling underlying marginalization through situationally sensitive actions.

Long-term sustainability

Driving sustainability requires agencies to transform how they conduct business at every level. Sustainability should be a part of developing an organizational strategy. Sustainability for organizations shows up in three areas, Environment, Social, and Governance (ESG). Sustainability is an important part of government services and programs. Sustainability is important to customers, employees and interested parties that recognize sustainability as an indicator of a program/projects long-term viability.

Adopting sustainability in the long term helps enterprises stay ahead of shifts while also offering a way to become more innovative and efficient.

Questions to consider:

- Will the benefits of GenAI be sustainable for vulnerable communities over time?
- Are there plans for ongoing support and resources?
- What is the agency plan if sustainability is not possible?

Potential actions to take:

- Define sustainability goals.
- Engage with interested parties from the beginning.
- Conduct a sustainability assessment.
- Develop a sustainability plan.

- Monitor sustainability per the sustainability plan.
- Practice continuous improvement.
- Use key performance indicators (KPIs) to assess progress.

Evaluation and monitoring

Evaluation should be a participatory process that allows all interested parties to provide input into the effectiveness, benefits, and concerns about a system. Monitoring is an ongoing internal process that regularly audits the AI system's outputs for biased, incorrect, and/or hallucinated information.

Questions to consider:

- Is there a system for continuous evaluation of the impacts of generative AI on vulnerable communities?
- How will feedback from these communities inform future iterations of the technology?
- Does the agency have the resources needed to properly evaluate and monitor AI tools?

Potential actions to take:

- Use the risk assessment guidelines associated to the EO to develop methods for reviewing AI outputs.
- Become familiar with the common GenAI risks per the NIST GenAI Profile.
- Ensure that there is a clear mechanism for providing feedback on all public (and potentially private) service offerings.
- Develop a tool kit and training for how employees can identify and report incorrect outputs in AI systems.

Conclusion

Equitable access and outcomes related to vulnerable communities are an essential aspect for state agencies to understand when deploying GenAI technologies. This understanding ensures that everyone has access to the same opportunities and support but tailored to their specific needs and circumstances.

Effective strategies deployed by agencies including targeted resource allocation, inclusive policymaking, active community engagement, continuous assessment and improvement is one part of helping vulnerable communities overcome barriers. Addressing barriers such as systemic biases, resource constraints, and resistance to change requires commitment, collaboration, and innovative approaches.

By prioritizing equity in decision-making, agencies can achieve outcomes that matter. By acting with integrity, we promote fairness and lead to stronger, more resilient communities and organizations. By emphasizing the positive effects of diverse thoughts and experiences, we foster inclusion. Through collaborative partnerships that focus on these principles and actions, state governments and communities can work towards achieving equitable outcomes. Ultimately, we are creating a more just and inclusive environment where everyone can thrive.

References

Executive Order 24-01 on Artificial Intelligence Section 5. requiring this deliverable reads as follows:

By December 2024, WaTech, in consultation with the Office of Equity, community members, Tribal governments, subject matter experts, and other relevant state agencies, shall develop guidelines for agencies to analyze the impact that adopting generative AI technology may have on vulnerable communities, including criteria to evaluate equitable outcomes in deployment and implementation of high-risk use cases. These guidelines and criteria shall inform whether and how an agency deploys a particular AI System. WaTech shall periodically revise the guidelines and criteria, when appropriate, and establish a consultative process with academia, industry experts, and representatives from advocacy organizations that represent communities that are disproportionately vulnerable to being harmed by algorithmic bias. WaTech and the Office of Equity shall make the guidelines publicly available by posting on their respective agency's website.

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 Health Organization: [Vulnerability and Vulnerable Populations](#)
- WA State Department of Health: [Community Engagement Guide](#)
- King County: [Community Engagement for Preparedness Guide](#)
- WA State Office of Equity: [Framework for Accountability in Generative Artificial Intelligence for Washington State Agencies](#)
- WA State Office of Equity: [shared-power-principle.pdf](#)
- WA State Office of Equity: [EquityHubResources_Relational Partnership Guide.docx](#)
- Procurement Guidelines for GenAI
- [Washington State Agency Privacy Principles](#)

- [RCW 43.376 - Government-to-Government Relationship with Indian Tribes \(Full RCW as PDF\)](#)
- [Centennial Accord](#)
- [Out of State Accord](#)
- [Millennium Agreement](#)
- [Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification](#) Joy Buolamwini joyab@mit.edu MIT Media Lab 75 Amherst St. Cambridge, MA 02139. Timnit Gebru timnit.gebru@microsoft.com Microsoft Research 641 Avenue of the Americas, New York, NY 10011
- [\[2402.14875\] What's in a Name? Auditing Large Language Models for Race and Gender Bias \(arxiv.org\)](#)
- [ChatGPT is biased against resumes with credentials that imply a disability – but it can improve | UW News \(washington.edu\)](#)
- [AI image generator Stable Diffusion perpetuates racial and gendered stereotypes, study finds | UW News](#)
- [Exploring LGBTQ+ Bias in Generative AI Answers across Different Country and Religious Contexts: Lilla Vicsek, Anna Vancsó, Mike Zajko, Judit Takacs. Cornell University. July 2024](#)