2026 OREGON STATE
EQUITY DATA SUMMIT

DESIGNING EQUITY IMPACT & RISK ASSESSMENTS

*FOR **AI & NEW TECHNOLOGY***

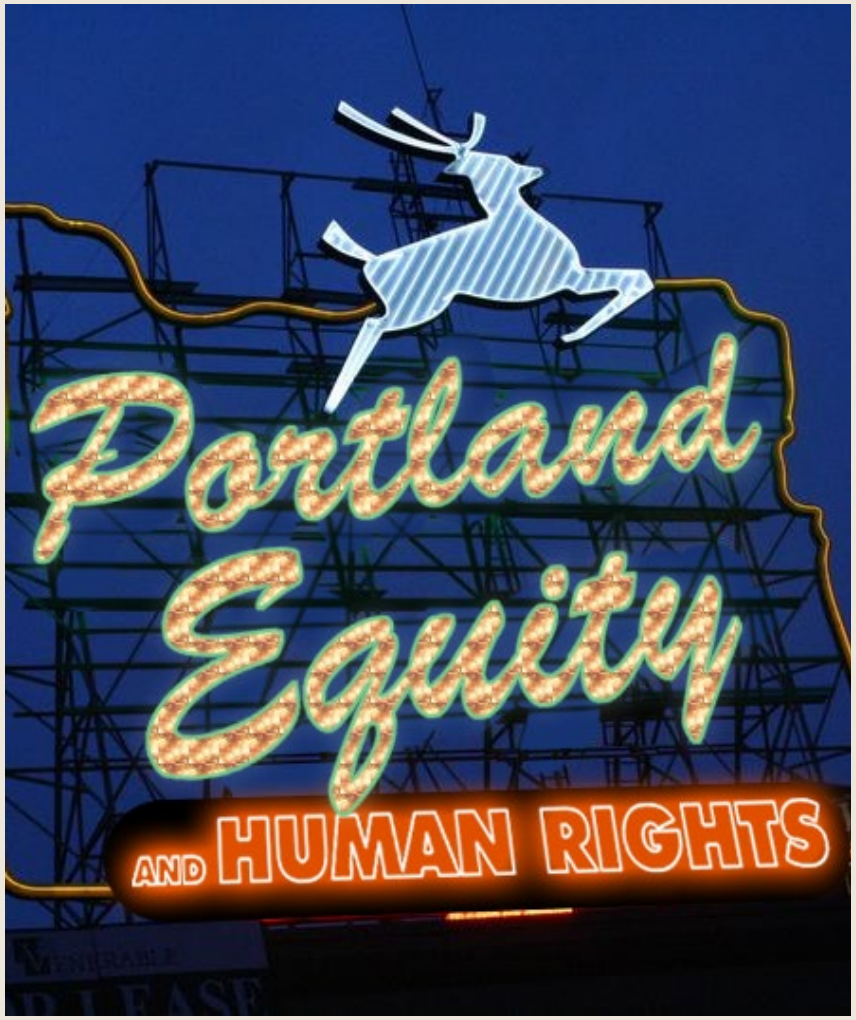


Aaron Garber-Paul

Equity Data and Policy Analyst

Office of Equity & Human Rights

City of Portland



CITY OF PORTLAND OFFICE OF EQUITY & HUMAN RIGHTS

The Office of Equity & Human Rights was created by Council Ordinance No 184880.

Our role is to coordinate the work of City bureaus in **helping the City to become a leader in defining and achieving equitable outcomes and eliminating disparities among Portlanders.**

We set the Equity vision and goals for the City of Portland

DESIGNING
EQUITY IMPACT & RISK ASSESSMENTS
*FOR **AI & NEW TECHNOLOGY***

1. **FOUNDATIONS**
2. **FRAMEWORKS**
3. **APPLIED PROCESSES**
4. **EQUITY BIAS & RISKS**

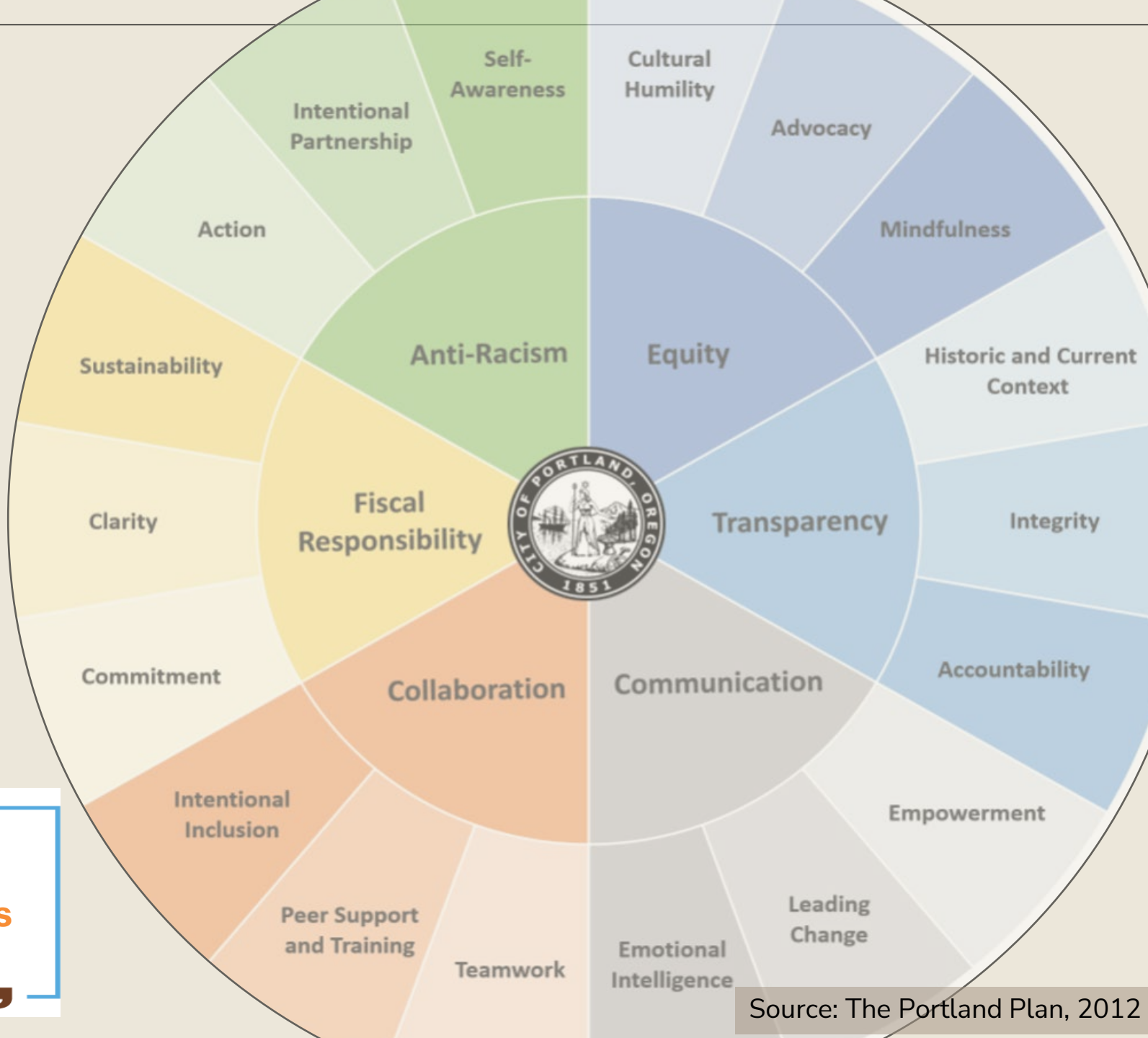
FOUNDATIONS.

EQUITY *(def)*

Equity (1): "When social identities do not determine or predict the distribution of City resources, opportunities, and burdens for individuals or communities."

Equity (2): "Equity is when everyone has access to the opportunities necessary to satisfy their essential needs, advance their well-being, and achieve their full potential."

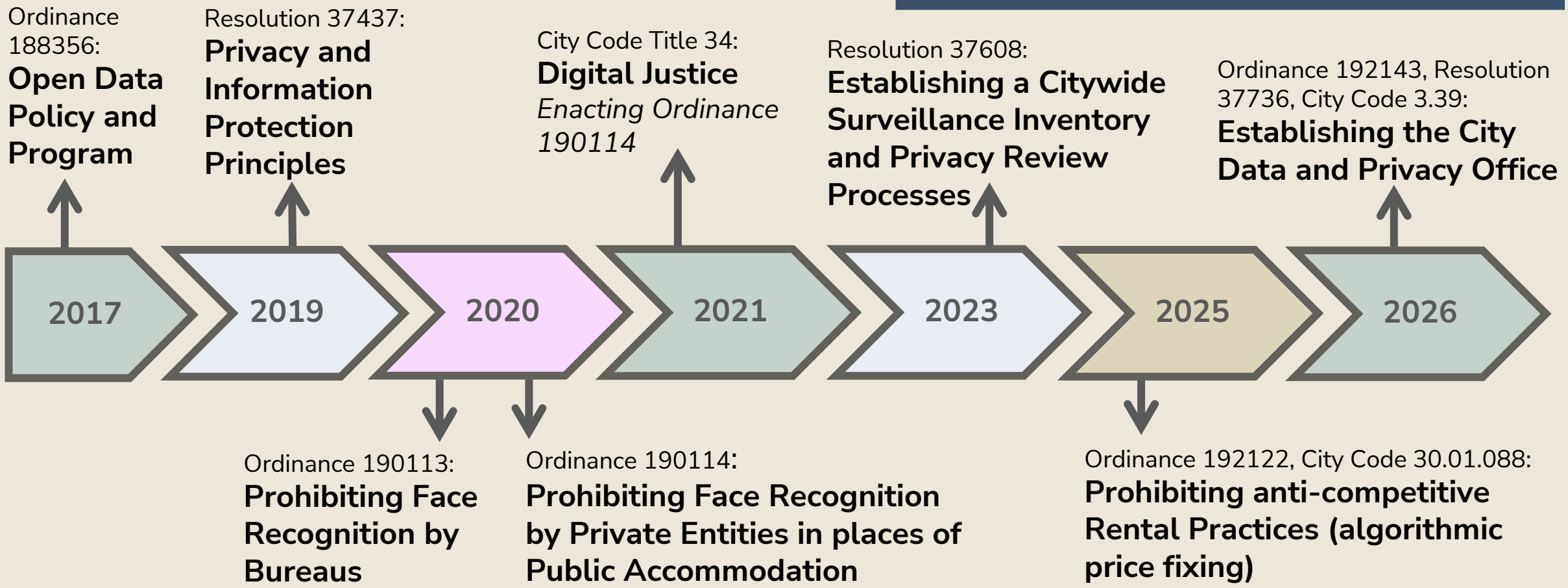
> **Practicing Equity is everyone's responsibility and is a continuous improvement process.**



Source: The Portland Plan, 2012

POLICY TIMELINE

> *"Privacy and Information protection principles center Equity and Human Rights in Privacy strategy development and acknowledge that underserved communities are most at risk in the digital age,"*
>
> *Resolution 190113 (2020)*
> *Banning Facial Recognition Technologies*



CONSUMER PRIVACY RIGHTS IN OREGON

The right to access personal data that has been collected about you.

The right to know a list of the specific entities that have received your personal data or any personal data from a business.

The right to correct inaccuracies in your personal data.

The right to have your personal data deleted.

The right to obtain a copy of your personal data.

The right to say “no” to (opt-out of) a business selling your personal data

FOUNDATIONS FOR EQUITY-CENTERED RISK ASSESSMENTS *OF* AI

City of Portland

- Equity Decision/ Data tools
- Privacy & Equity Impact Assessment (PIA) process
- Policies on AI, and AI Principles
- Guidance on Information Security, Retention, et all

State law

- Oregon Consumer Privacy Act
- Executive Order on AI *(new!)*

Federal law

- (next slide)

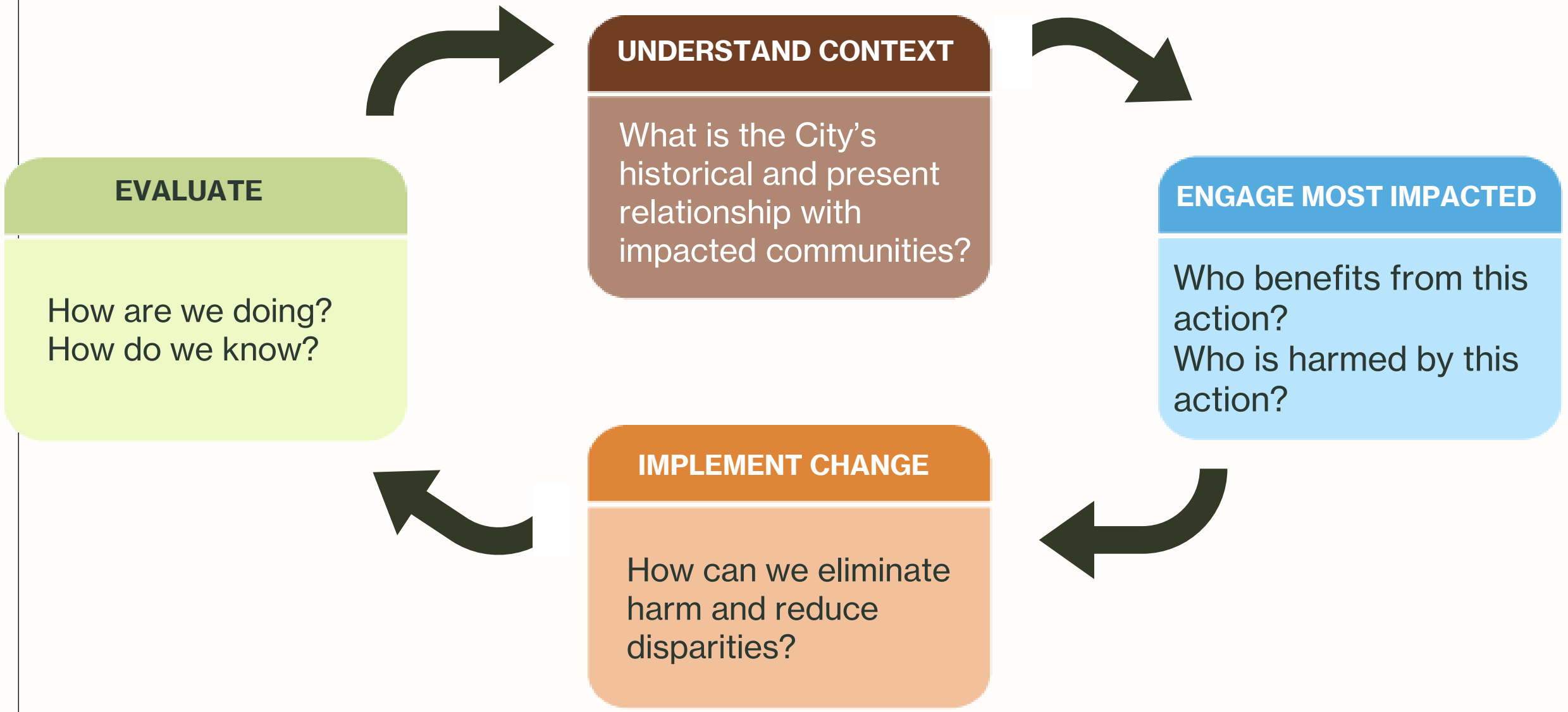
International and Professional guidance:

- NIST Gen AI Risk Management Framework, associated publications
- OECD AI Principles
- EU AI guidelines, UNESCO Recommendation on the Ethics of AI
- Social research *methods and other disciplines*

EQUITY & CIVIL RIGHTS COMPLIANCE

- **Civil Rights Title VI**
  - Non-Discrimination
  - Fairness & Equal Opportunity
  - Language Access
- **ADA Title II**
  - Accessibility
  - Equal service
- **LGBTQ2IA**
- **Gender Equity**
- **Environmental Justice**
- **Religious protections**
- **Free Speech/ Political Opinions & groups**

Equity Evaluations *are Iterative:*



Trustworthy and Responsible AI



"Trustworthy and Responsible AI is not just about whether a given AI system is biased, fair or ethical, but whether it does what is claimed."

- NIST 1270 (2022)

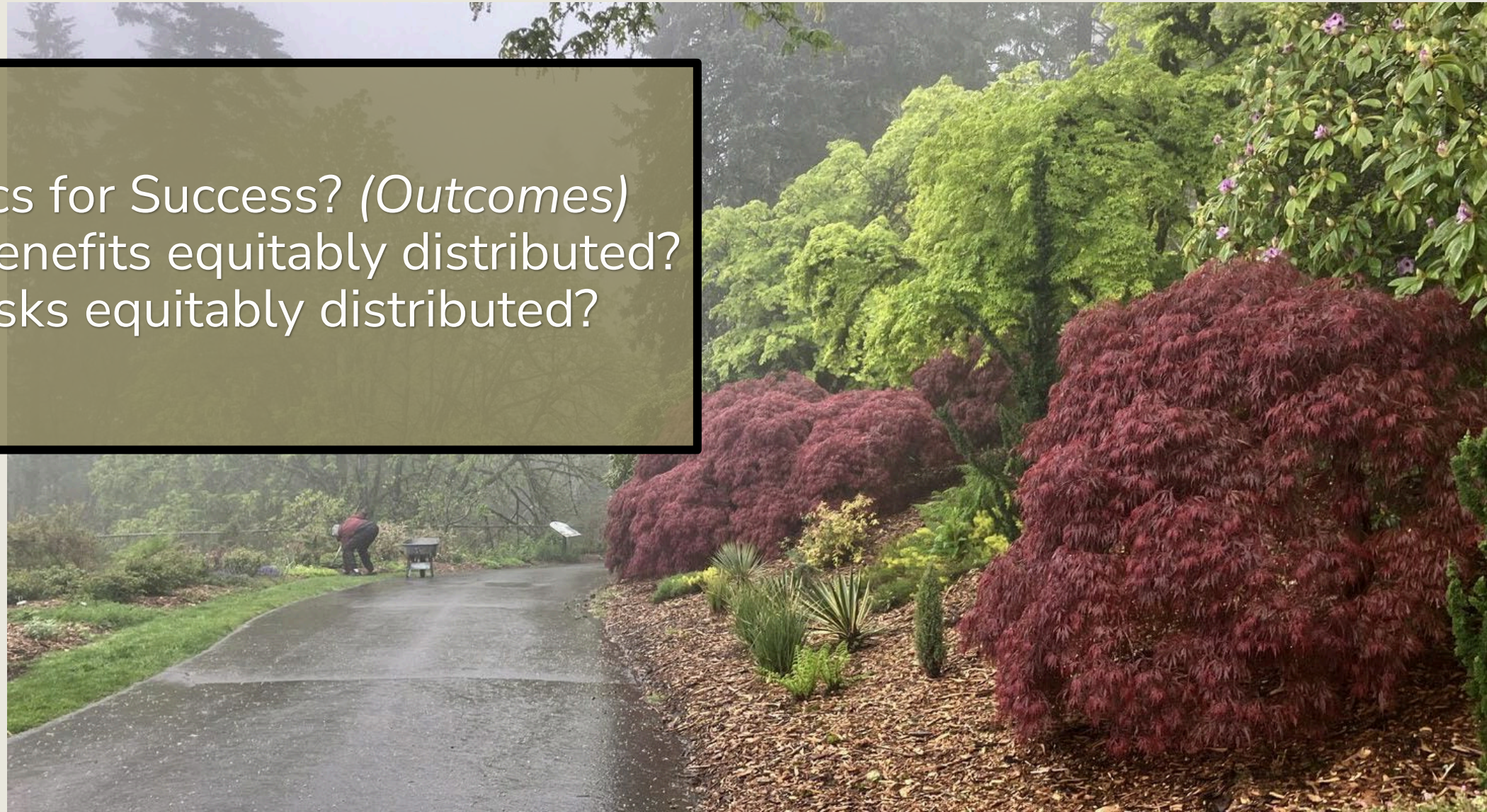
For each prediction, we want to know:

- ❑ Is this output correct?
- ❑ How confident are we in that judgement?
- ❑ Why is it correct or incorrect?
- ❑ How would we score the quality?

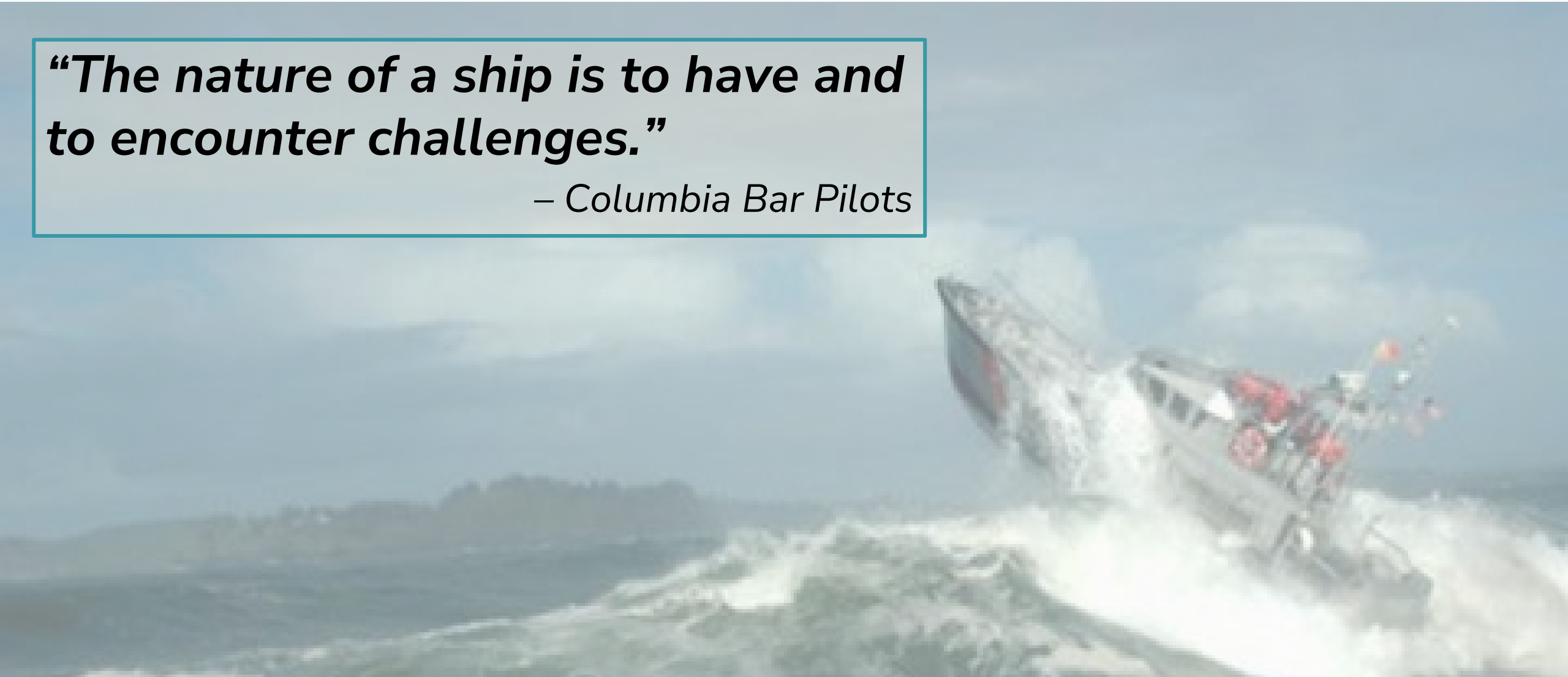
ASSESSMENT FRAMEWORKS

1. How well do we succeed?

- Metrics for Success? *(Outcomes)*
- Are benefits equitably distributed?
- Are risks equitably distributed?



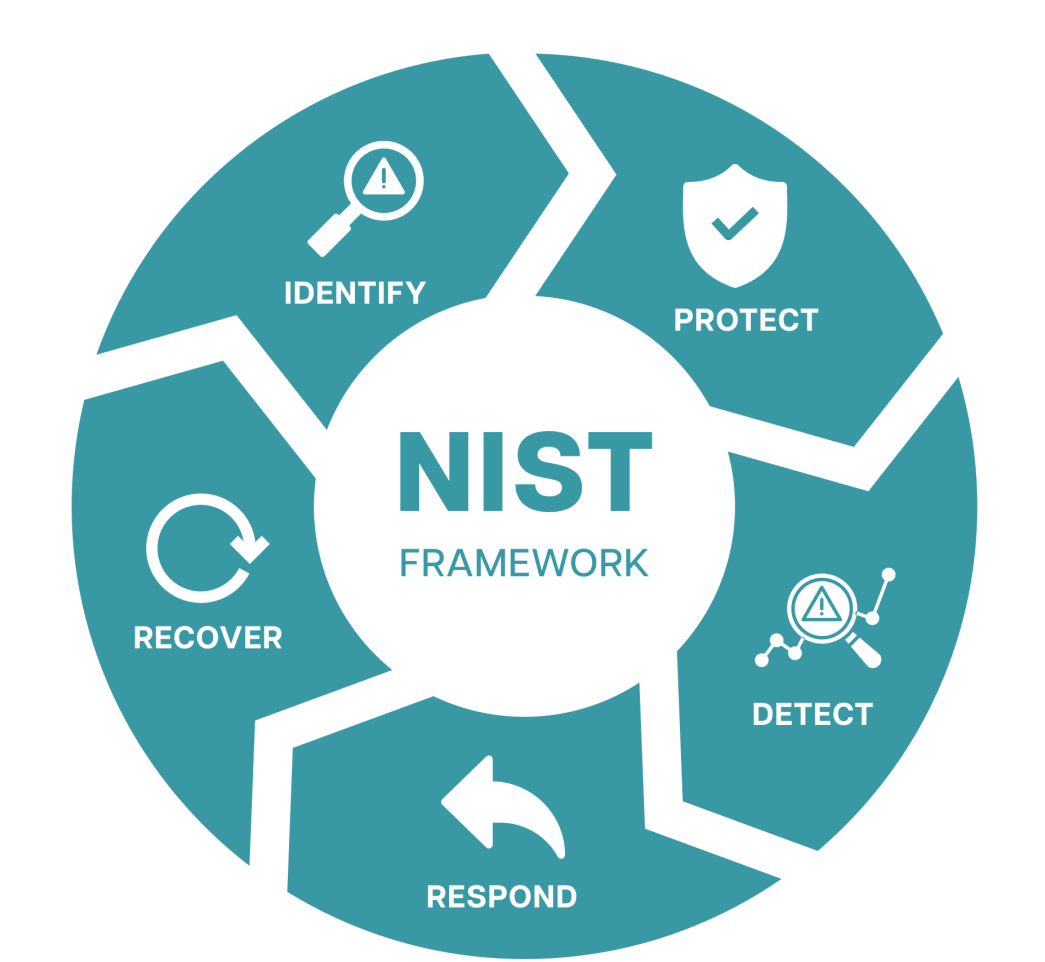
2. How well do we fail?



"The nature of a ship is to have and to encounter challenges."

– Columbia Bar Pilots

AI RISK - BEST PRACTICES



NIST Incident Response Planning

Guiding Standards:

- Data minimization
- Explainability, Transparency
- Auditability, audit logging
- **Humans in control**
- Quality control, accuracy planning
- Routine Bias-testing
- Appropriate access controls, including role-based security, logging access.
- Data retention, lifecycle planning, public records planning

- ❑ **Complaint process**
- ❑ **Incident response plan**
- ❑ **Third party audit**

AI RISK - BEST PRACTICES

“HUMAN IN CONTROL”

- Human has real authority
- Can stop, change, override
- Understands what the system does and why
- Sets the pace for critical choices
- Retains responsibility and insight

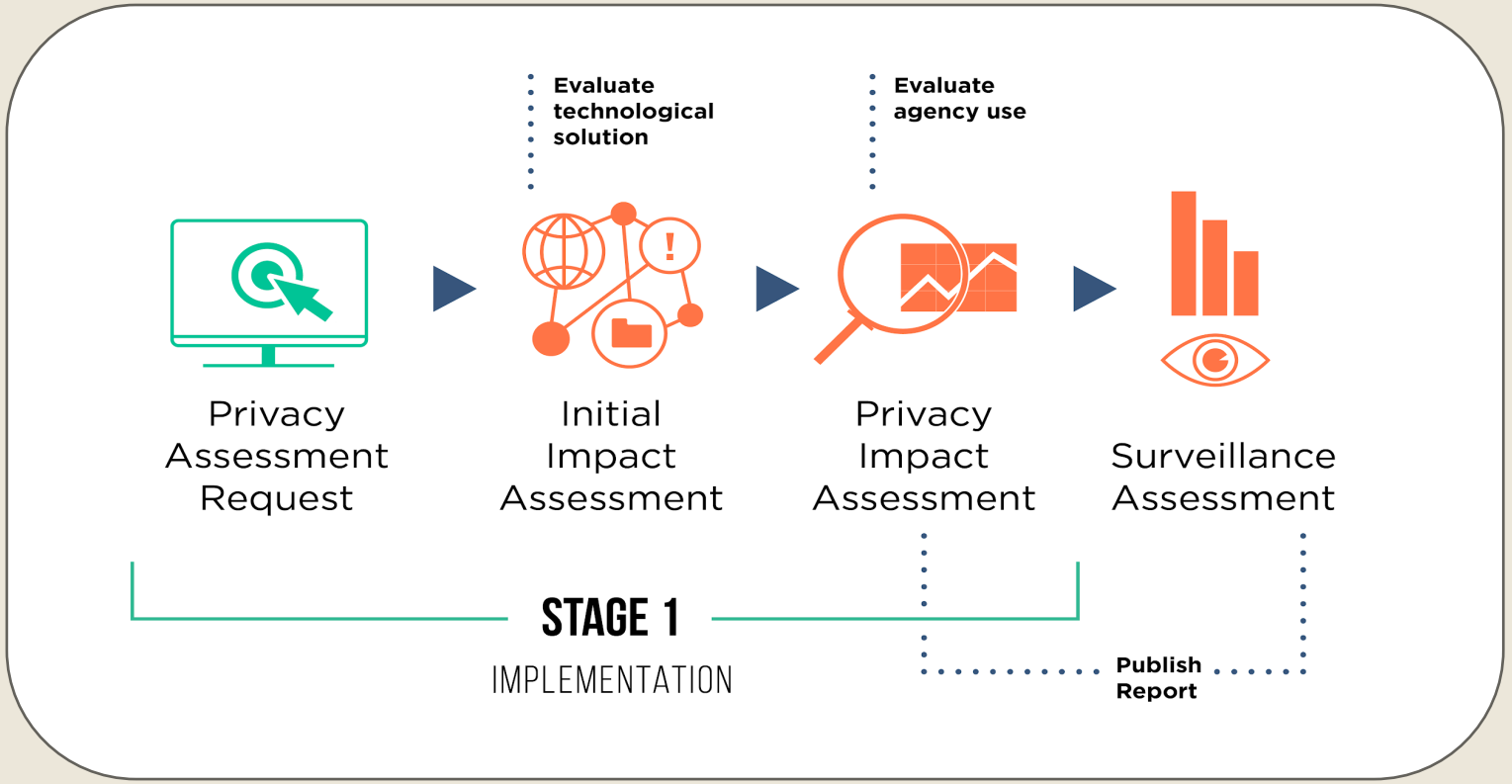
- **Efficacy**
- **Curiosity**
- **Accountability**
- **Empathy**

CITY OF PORTLAND

APPLIED PROCESS

CITY OF PORTLAND PRIVACY IMPACT ASSESSMENTS (PIA)

WORKFLOW



Intake	Research and Analysis	Internal Reviews	Bureau comments	Integration and final
Intake form is filled out by requesting bureau.	• Research on technology, use, and context. • **First draft is created.**	• PIA draft is reviewed by Legal, Equity, Privacy, and technology teams • **Second draft is created**	Bureau reviews second draft and provide comments for accuracy and refinement.	• **Final draft.** • **Released version.**

PRIVACY IMPACT ASSESSMENTS

RISK MATRIX

SEVERITY (BY RISK)

PROBABILITY

	Low	Moderate	High	Extreme
Almost Certain				High
Likely				
Possible		Medium		
Unlikely				
Rare	Low			

RISK TYPES

Individual Privacy Harms	Equity, Disparate Community Impact	Political, Reputation & Image	City Business, Quality & Infrastructure	Legal & Regulatory	Financial Impact

PROBABILITY ASSESSMENT

Likelihood	Probability
Almost certain	Likely to occur yearly
Likely	Likely to occur every 2 years
Possible	Likely to occur every 5 years
Unlikely	Likely to occur every 10-20 years
Rare	Has never occurred

Privacy Impact Assessments: **Evaluation**

	Individual Privacy Harms	Equity, Disparate Community Impact	Political, Reputation & Image	City Business, Quality & Infrastructure	Legal & Regulatory	Financial Impact
low	Customer or "telephone book" information collected and could be disclosed (excluding utility customer data, protected by RCW)	**Little or no equity impact, technology delivered uniformly without reference to individuals or demographic groups**	Issues could be resolved internally by day-to-day processes; little or no outside stakeholder interest.	Management of disclosure issues would represent negligible business interruption; resolved with no loss of productivity	Adverse regulatory or legal action not indicated or highly unlikely	$0-$500 impact; internal costs covered, and no significant external costs incurred
medium	Potential disclosure would be limited to non-financial, non-health related information; no personal identifiers (e.g., social security and driver's license #s)	**Accidental or perceived disparate impact to communities by nature of location of technology or service delivered**	Issues could be raised by media and activist community resulting in protests and direct community complaints	Issue management would result in brief loss of services; loss of < 1 week service delivery; limited loss of productivity	Relatively minor incident, regulatory action unlikely; possible legal intervention or consultation for addressing data exposure or loss	>$500 - $5,000; internal and external costs associated with legal consultation, system rework, overtime
high	Financial or other highly sensitive information would be collected and disclosable requiring action to remediate negative effects (example: non-HIPAA health data); i.e., credit report management required	**Intentional disparate equity impact resulting in community concern resulting in privacy harms, media coverage; loss of reputation, legitimacy and trust impacted**	Disclosure would likely result in heavy local media coverage; reputation, legitimacy and trust impacted	Significant event; loss of > 1–3-week loss of services; critical service interruption to delivery of infrastructure services	Adverse regulatory action likely – i.e., fines and actions associated with CJIS, HIPAA, PCI, NERC, COPPA violations, etc.	> $5,000 -$50,000 external costs associated with fines, consultation fees and regulatory actions to mitigate information exposure; internal costs associated with system rework, overtime
extreme	Disclosure would result in extreme privacy impacts on highly regulated information; catastrophic public release of financial and personal information requiring credit report monitoring and other remediation	**Extreme impacts to community, City experiences national media attention; widespread public concern and protest; significant breakdown in business processes associated with damage control**	Likely national and international media coverage; serious public outcry; significant breakdown in business processes associated with mitigation and damage control	Extreme event; business collapse for department services; loss of > = 3 months of data or productivity; critical business infrastructure loss > 1 month	Major legislative or regulatory breach; investigation, fines, and prosecution likely; class action or other legal action	> $50,000 external costs associated with fines, consultation fees and regulatory actions to mitigate information exposure; internal costs associated with system rework, overtime

PRIVACY IMPACT ASSESSMENTS

EXAMPLE:

FINAL DRAFT

2025_06_30 BPS- RECYCLING CONTAMINATION DETECTION TOOL

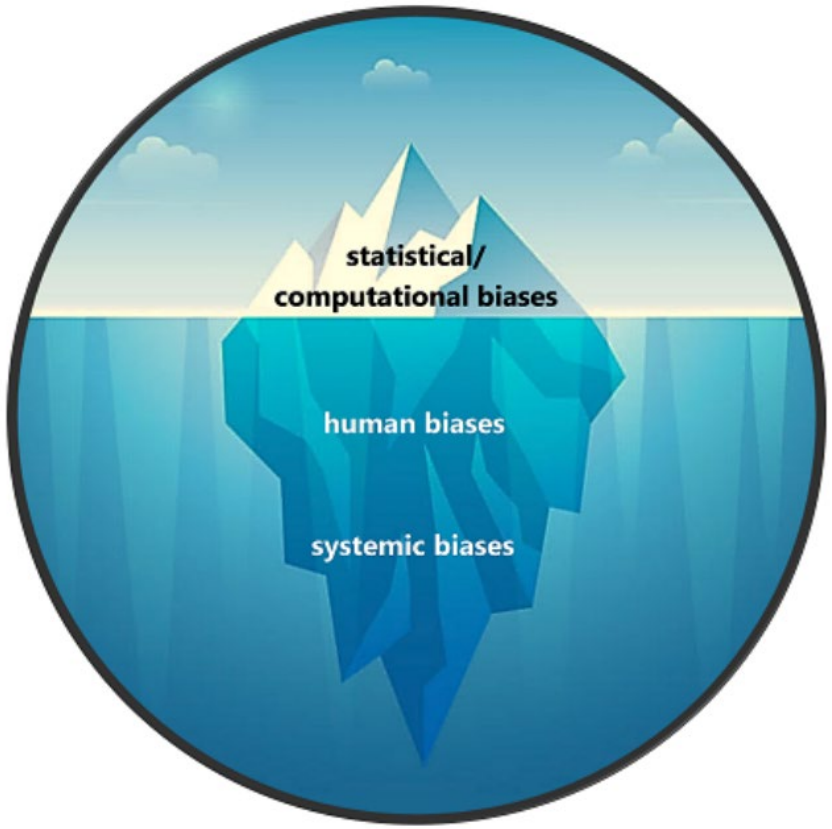
Privacy Impact Risk Severity Assessment

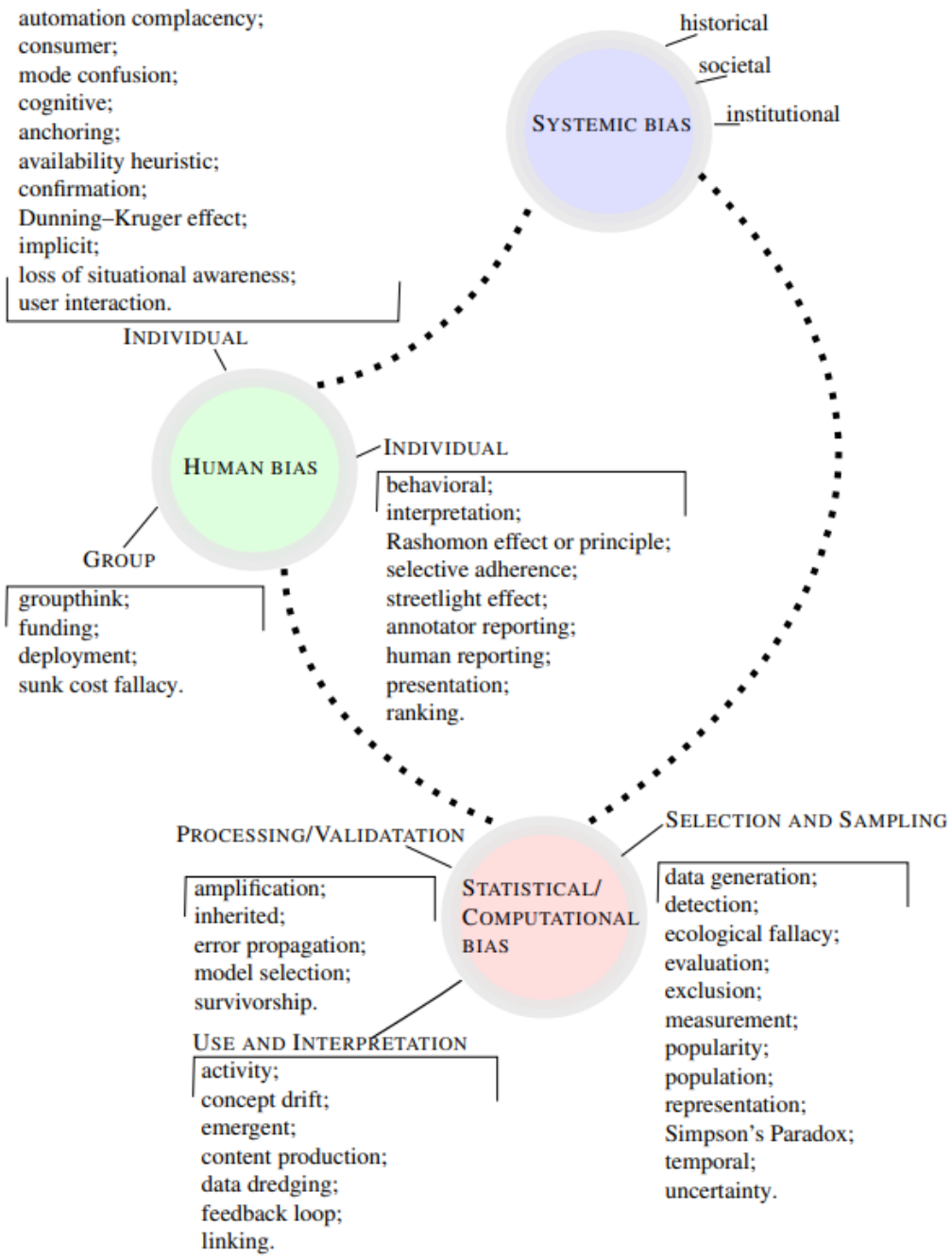
WORST CASE SCENARIO	Medium

Baseline (B): (T) – Technology level, (U) – use and application level. Risk type (RT): (I) Individual Privacy Harms; (II) Equity, Disparate Community Impact; (III) Political, Reputation & Image; (IV) City Business, Quality & Infrastructure; (V) Legal & Regulatory; and (VI) Financial Impact.

B	RT	Risk description	Impact	Likelihood	Mitigation, comments, and strategies	Risk level
T	II	2.1 Risk of bias in training data for identifying recycling items or bias in the image detection algorithm	Moderate	Possible	Factors that make up for this risk are mainly from the training data set and the training the machine learning model goes through before it is deployed. If the system is better at identifying unauthorized objects from minorities' recycling bins than unauthorized objects from another groups (or vice versa), then the City may give fines or notices inequitably. A process of constant improvement implemented by the technology vendor is recommended.	Medium
T	II	2.2 Risk from unequal use of technology due to companies' uneven implementation.	High	Possible	Haulers may use different vendors and type of cameras (e.g., quality, brand, etc.). Because they cover different parts of Portland, it can create differences between neighborhoods recycling collection services could create disparities in how actions from contamination detection are implemented. Implementation of technology needs to fulfill a baseline or standards defining quality of services and privacy and cybersecurity safeguards. Collection of metrics that report back technology effectiveness (false positives, identifying the type of contamination), contamination detection aggregated by geography. Releasing a dashboard accessible publicly with information about contamination identification and false positives or other relevant issues can create visibility on some potential issues derived from uneven quality of vendors' technology solutions. This risk is reduced if the City has control over the technology offered by vendors.	Medium

IDENTIFYING RISK & BIAS IN AI





Q: Where does Equity Bias in AI come from?

Tools learn from our past decisions.

If the past treated people unequally, the tool learns that as the pattern and applies it to every new case.



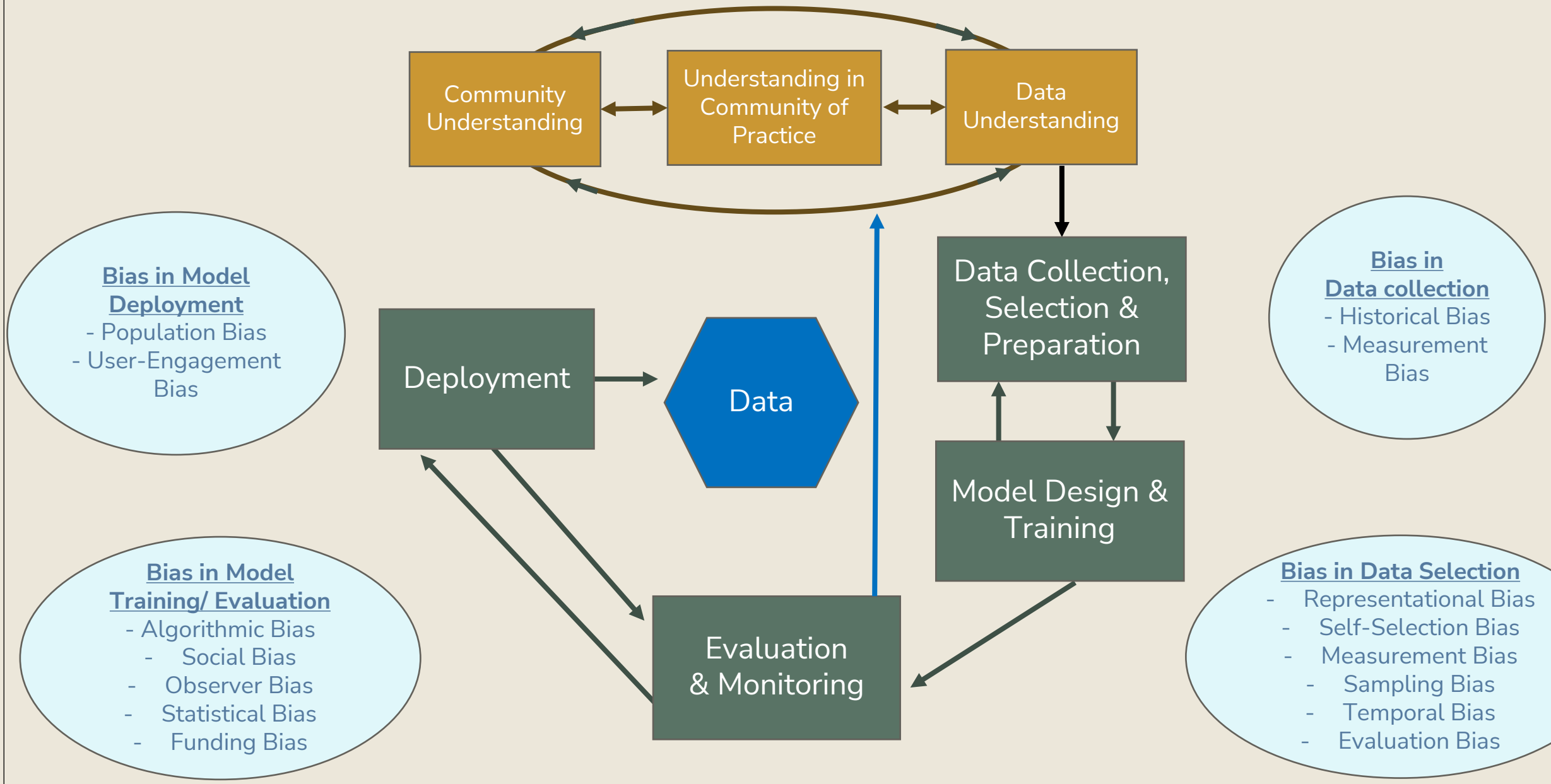
IMPLICIT BIAS

A prejudice or stereotype that is present but not consciously recognized.

SELECTION BIAS

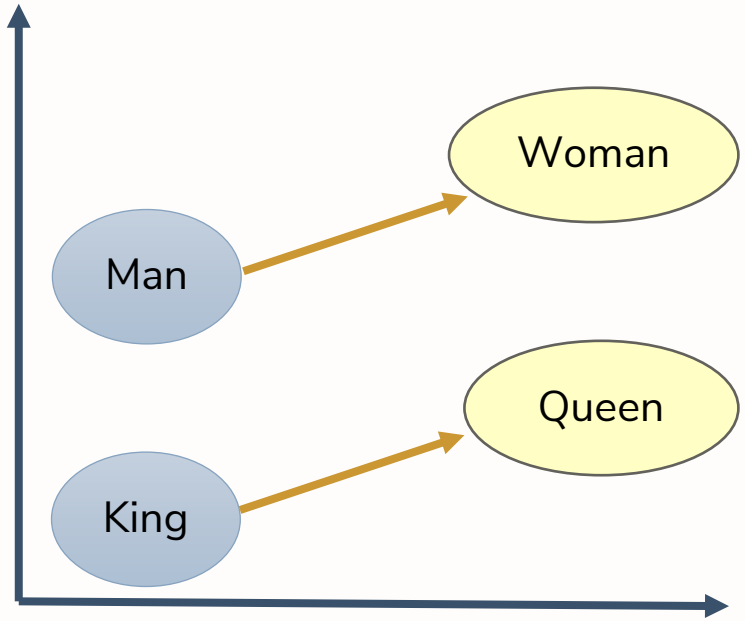
A logical error which occurs when analysis or modeling does not consider the people or data that did not make it past a particular selection process and into the dataset.

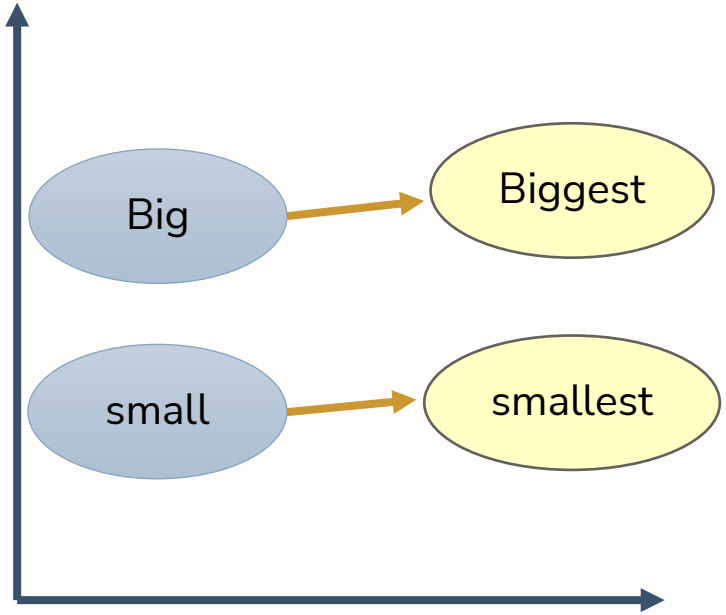
EQUITY BIAS IN AI DEVELOPMENT PIPELINES:

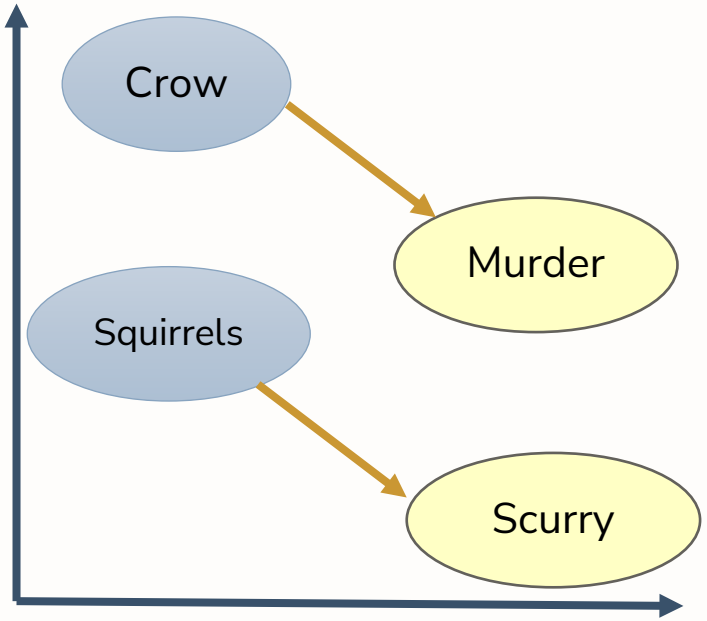


Automating Equity Bias

Vectors are the language of neural networks, built from the way humans use words.







Does *Doctor – Man + Woman = Nurse*?

AI & Language Access Risks

<u>Equity requires more than accurate translations</u>:

If someone asks an AI model the same questions in two different languages, will they get the same answer?

<u>No</u> City Staff are certified in written translation

The Citywide Language Access Program develops policy, standard operating procedures, tools, resources, and guidance based on national and federal standards and best practices. The program seeks to increase equitable access for community members who experience institutional language barriers

Q: Is AI the right solution?

FALLACY OF OBJECTIVITY

TECHNO-SOLUTIONISM

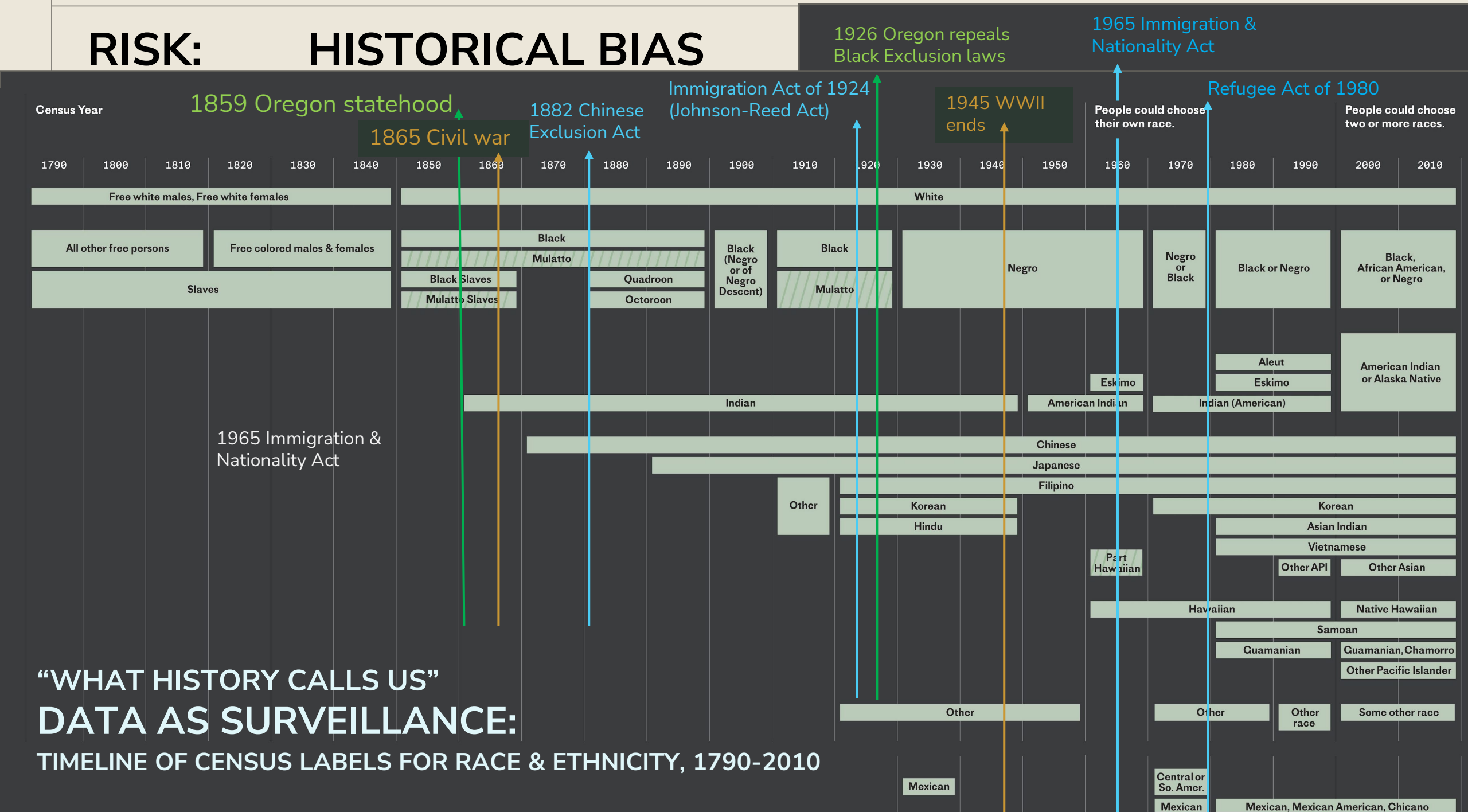
McNAMARA FALLACY

TECHNO-CHAUVANISM



Image generated with AI

RISK: HISTORICAL BIAS



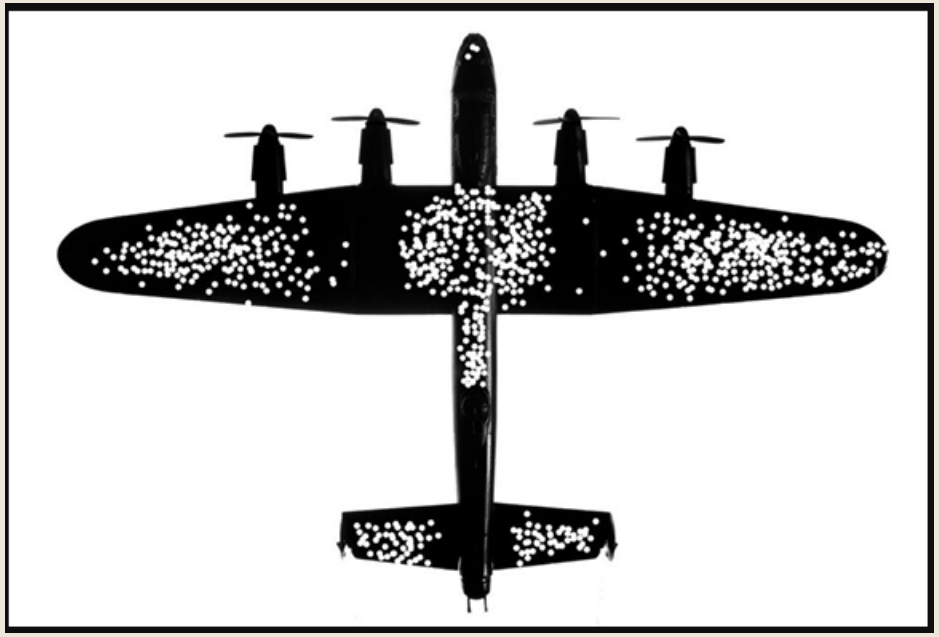
"WHAT HISTORY CALLS US"

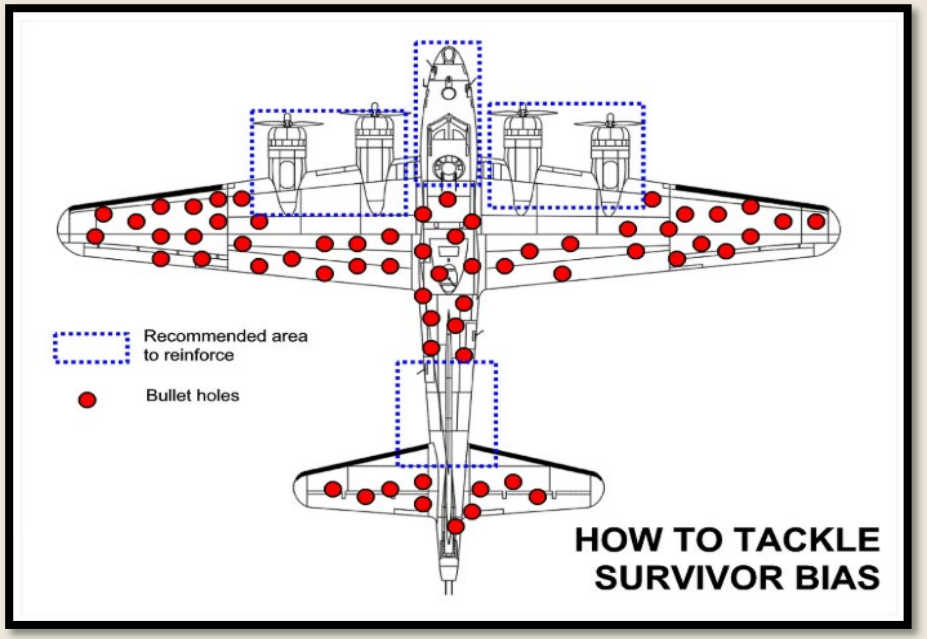
DATA AS SURVEILLANCE:

TIMELINE OF CENSUS LABELS FOR RACE & ETHNICITY, 1790-2010

SELECTION BIAS

A logical error which occurs when analysis or modeling does not consider the people or data that did not make it past a particular selection process and into the dataset.





REPRESENTATION BIAS

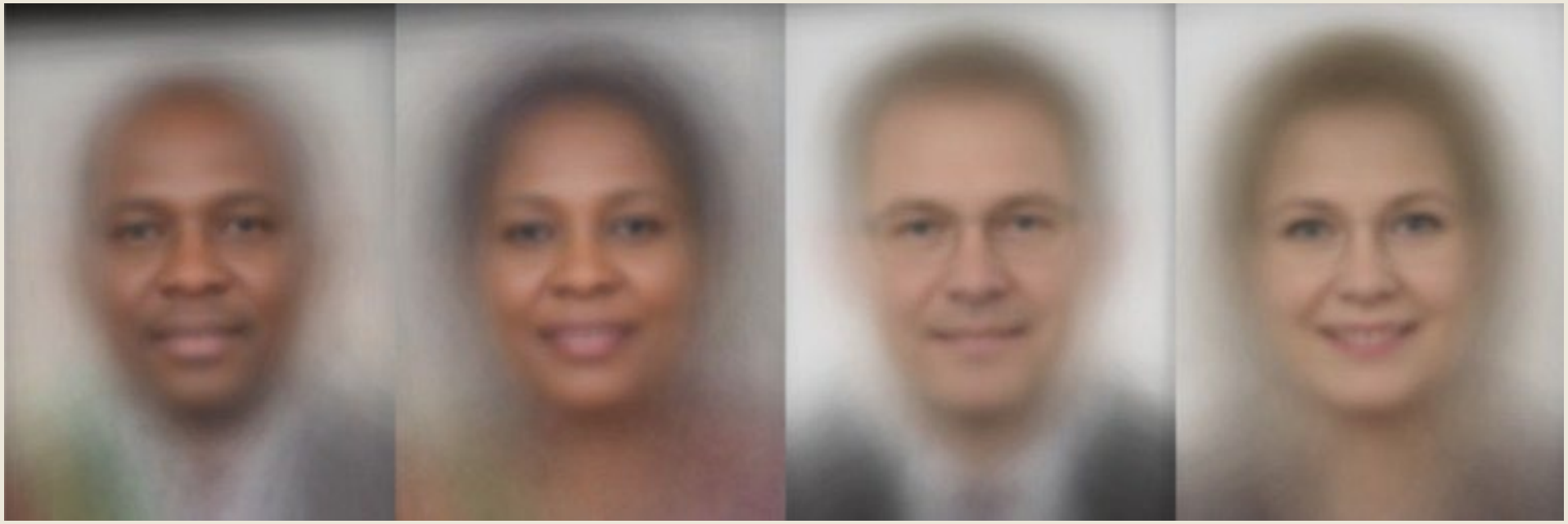


Figure 1 | A gender-balanced facial image bank with a range of skin tones
Credit: Dr Joy Buolamwini (CC BY 4.0)

ON “LEADERSHIP”





Images generated with AI

ON “UGLINESS”







Images generated with AI

RISK: OVERSURVEILLANCE

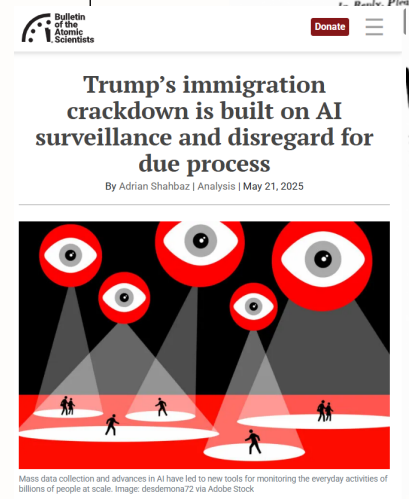
THE NEW PANOPTICON



Source: Doug Dubois & Jim Goldberg NYTimes 9-22-2002

Bias & disproportionate experiences of surveillance: Examples





~~SECRET~~

UNITED STATES DEPARTMENT OF JUSTICE
FEDERAL BUREAU OF INVESTIGATION

WASHINGTON, D.C. 20535

November 12, 1964

MARTIN LUTHER KING, JR.

1653

Confidential sources of this Bureau who have furnished reliable information in the past have supplied the following information concerning plans being made by Martin Luther King, Jr., and his associates concerning King's receipt of the Nobel Peace Prize which is to take place December 10, 1964, at Oslo, Norway.

On November 3, 1964, Bayard Rustin contacted the secretary of Dr. Ralph Bunche, Undersecretary for Special Political Affairs, United Nations (UN). Rustin stated that he was speaking on behalf of Martin Luther King, Jr., and requested to see Bunche regarding the receipt of the Nobel Peace Prize by King.

Subsequently on November 3, 1964, Rustin conferred with Bunche and told him that upon King's return to New York City, after receiving the Nobel Prize, there is to be a welcoming affair at either a large dinner at a hotel or an activity at Carnegie Hall. Rustin told Bunche that King wanted to know whether Bunche would be prepared to be chairman of the welcoming committee which will include civil rights leaders and "a number of key liberals." Bunche said that he would be glad to be chairman, but that he would have to be informed as to the others who will be on the

[~~CONFIDENTIAL~~ COMMITTEE PRINT]

81ST CONGRESS, 2d Session — SENATE — {DOCUMENT No. —

EMPLOYMENT OF HOMOSEXUALS AND OTHER SEX PERVERTS IN GOVERNMENT

INTERIM REPORT
SUBMITTED TO THE
COMMITTEE ON EXPENDITURES IN THE EXECUTIVE DEPARTMENTS
BY ITS
SUBCOMMITTEE ON INVESTIGATIONS
PURSUANT TO
S. Res. 280
(81st Congress)
A RESOLUTION AUTHORIZING THE COMMITTEE ON EXPENDITURES IN THE EXECUTIVE DEPARTMENTS TO CARRY OUT CERTAIN DUTIES

1950.—Ordered to be printed

118TH CONGRESS
1ST SESSION

H. R. 1064

To provide requirements related to the eligibility of transgender individuals from serving in the Armed Forces.

IN THE HOUSE OF REPRESENTATIVES
FEBRUARY 17, 2023
Mr. BANKS introduced the following bill; which was referred to the Committee on Armed Services

A BILL

To provide requirements related to the eligibility of transgender individuals from serving in the Armed Forces.

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled,

SECTION 1. SHORT TITLE.

This Act may be cited as the "Ensuring Military Readiness Act of 2023".

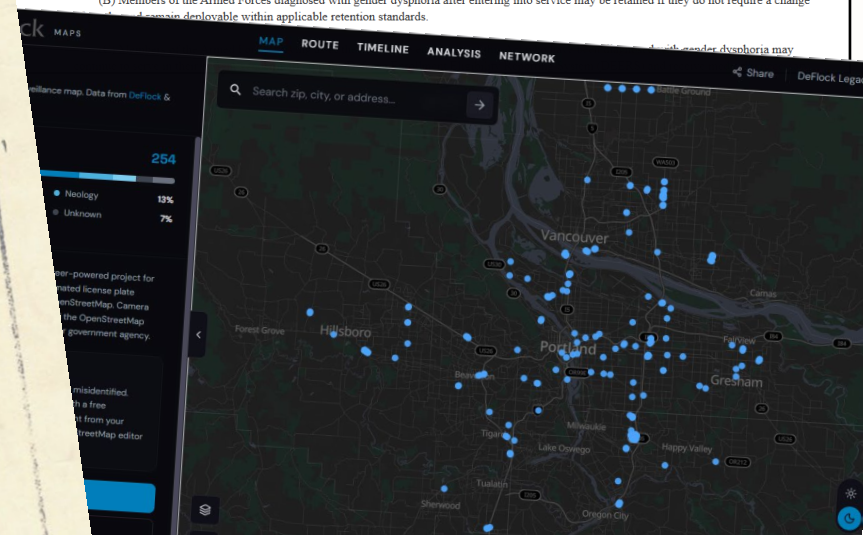
SEC. 2. LIMITATIONS ON MILITARY SERVICE BY TRANSGENDER INDIVIDUALS.

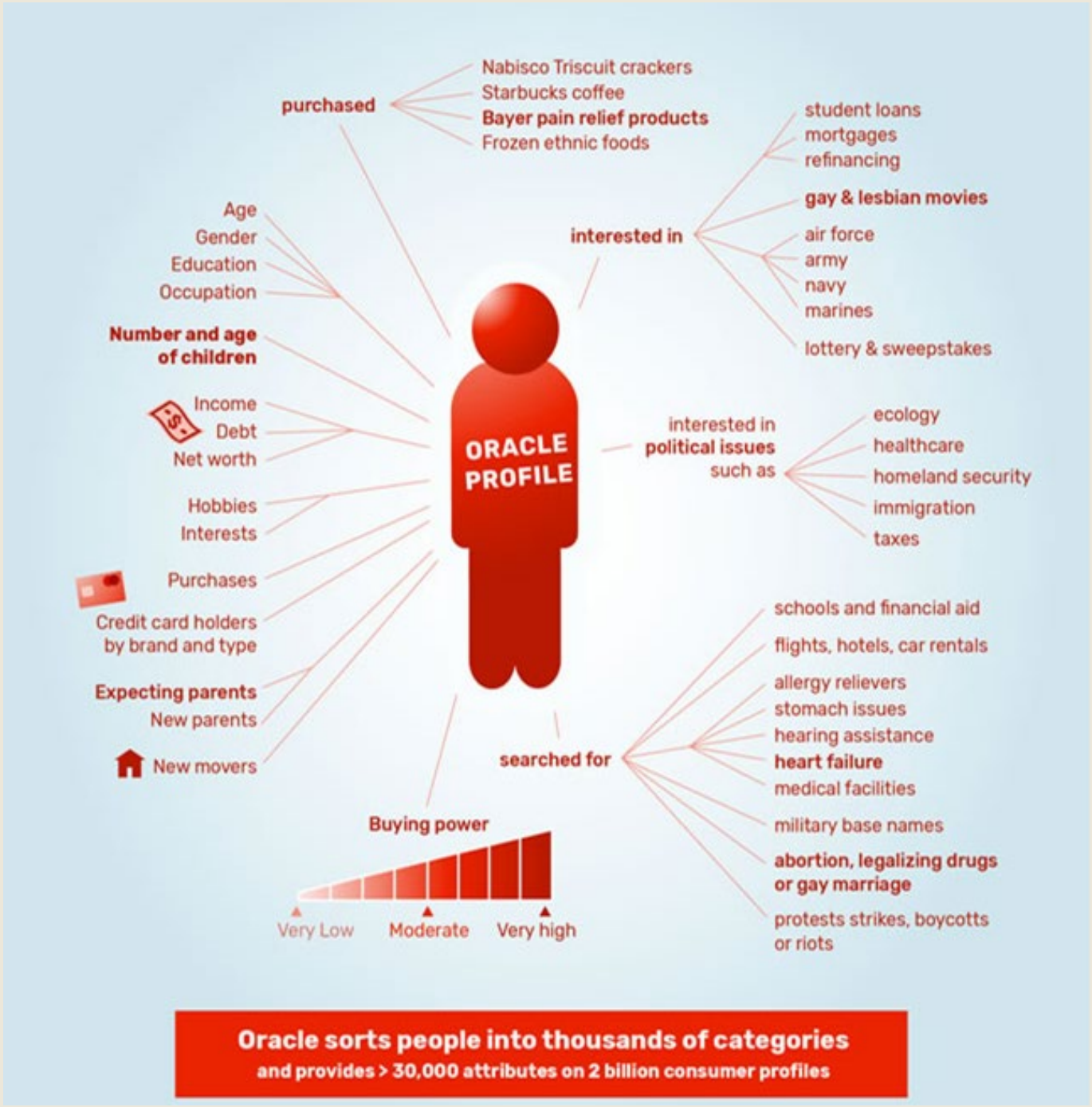
Not later than 90 days after the date of the enactment of this Act, the Secretary of Defense shall prescribe regulations regarding transgender service as follows:

(1) Transgender persons with a history of diagnosis of gender dysphoria are disqualified from military service except under the following limited circumstances:

(A) Individuals may serve in the Armed Forces if they have been stable for 36 consecutive months in their biological sex prior to accession.

(B) Members of the Armed Forces diagnosed with gender dysphoria after entering into service may be retained if they do not require a change





DATA PRIVACY RISK:

"Potential adverse consequences to individuals and society arising from the processing of personal data"

Community and individual risks from Reduced community connection points

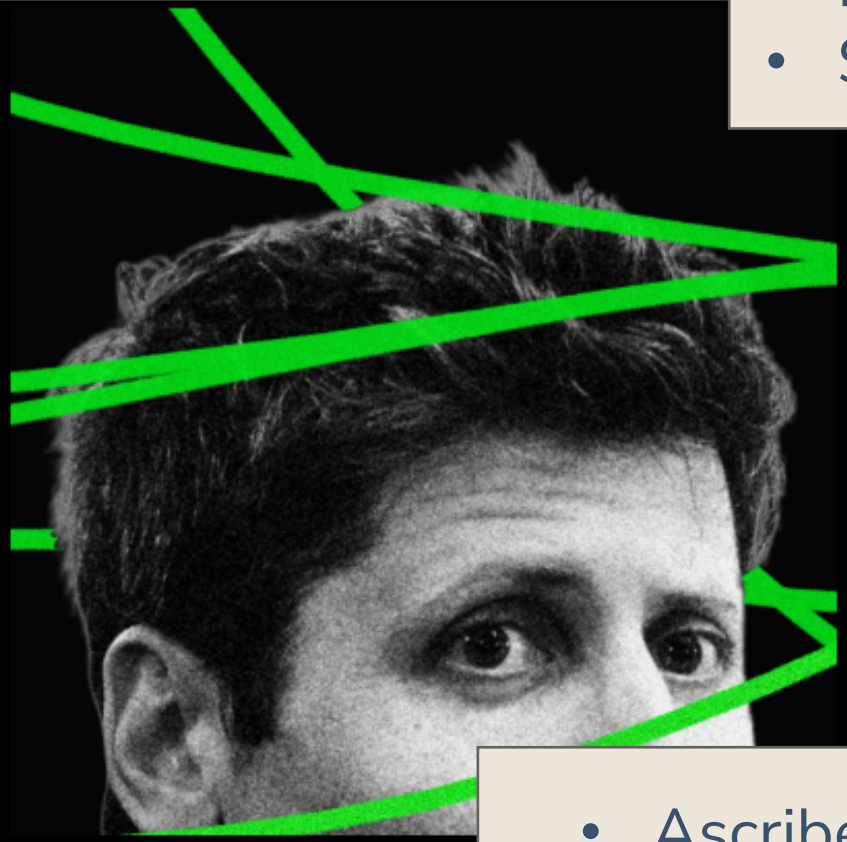
WIRED

NEWSLETTERS

LOUISE MATSAKIS BUSINESS OCT 27, 2025 1:00 PM

OpenAI Says Hundreds of Thousands of ChatGPT Users May Show Signs of Manic or Psychotic Crisis Every Week

OpenAI released initial estimates about the share of users who may be experiencing symptoms like delusional thinking, mania, or suicidal ideation, and says it has tweaked GPT-5 to respond more effectively.



- Mis-Information
- Echo Chambers
- Social Media Addiction

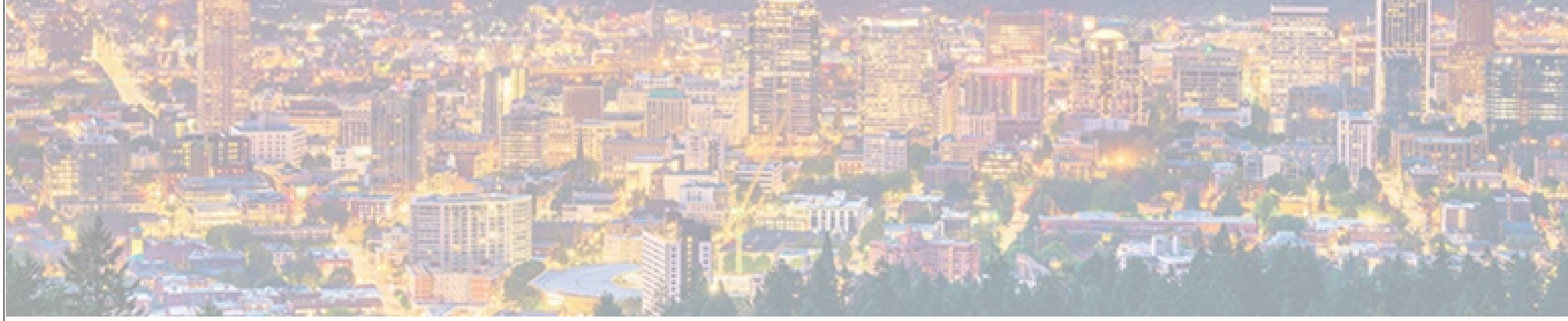
- Alienation
- Isolation
- Loneliness

- Suicidality

- Ascribed Consciousness

"Taxi drivers do more than drive"

- Office of Equity guidance on risks associated with autonomous, driverless taxis (2026)



WATCH THIS SPACE....

- Red light cameras
- **Facial recognition**
- Geolocation services
- **AI Translation**
- Smart meters
- Social networks monitoring
- Digital kiosks
- **Drones**

- **Automated license plate readers**
- Body worn cameras
- CCTV
- Connected transportation
- **Bluetooth and smart-body worn tech**
- Gunshot detectors
- **Automatic Professional Transcription (AI “note takers”)**