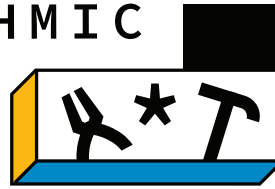


# ALGORITHMIC EQUITY TOOLKIT



## WHAT IS THE PROBLEM?

Hundreds of companies, big and small, are developing and marketing surveillance and decision-making technologies for government use. Technology can change government practices, and not always for the better.

## WHAT IS THE AEKit?

The Algorithmic Equity Toolkit (AEKit for short) is a collection of four components designed to identify surveillance and decision-making technologies used by governments; make sense of how those technologies work; and pose questions about their impacts, effectiveness, and oversight.

## HOW WAS THE AEKit CREATED?

The team behind the AEKit began working in 2019 to build the capacity of community groups to analyze government technologies. The AEKit is the result of a year of collaborative work between our team and the Tech Fairness Coalition, a collective of community groups convened by the ACLU of Washington.

## HOW DO I USE THE AEKit?

There are four components of the AEKit. They focus on identifying and analyzing technologies that aim to automate, aid, or replace human decision-making. The components of the AEKit are intended to be used in the order listed below, but feel free to use them as you see fit.

1

### FLOW→CHART

A chart with yes/no questions to help identify whether a technology is an automated decision system or a surveillance tool, both, or neither.

2

### SYSTEM MAP

A map that defines key terms used in the AEKit and links the relationships between the various parts of automated decision systems.

3

### FILL-IN-THE BLANK

A set of open-ended prompts that help to explore potential impacts of these technologies.

4

### QUESTIONNAIRE

A set of questions about automated decision systems to ask government employees, elected officials, and vendors.

This page has been intentionally left blank. Please print me double-sided.

# FLOW→CHART

Use this flowchart to identify whether a particular technology is an **automated decision system**. On the next page you can find definitions for the **bolded words**, and a map of the relationships between various parts of an **automated decision system**.

## ALGORITHMIC EQUITY TOOLKIT



*Automated decision systems pose certain hidden risks because of their use of data and algorithms. Identification of automated decision systems can be an important first step to intervening in the use of these systems.*

### START

The technology I am assessing is called:

.....

Does the technology make a record of, or do something in response to **input data**?  
*(For example: does it respond to words, photos, sounds, videos, clicks, or location data?)*

YES

NO

Does the technology make or help people make guesses, predictions, or suggestions?  
*(For example: does it create gender or race labels from a photo of a person's face, or make a suggestion about where future policing should focus based on crime statistics)*

YES

NO/NOT SURE

Does it use other **recorded data**?  
*(For example: does it use databases, maps, government statistics, laws and ordinances, or social media profiles?)*

YES

NO

Does the technology...

- make annotations to...
- find patterns in...
- visualize...
- draw connections within...
- automatically make changes to...
- identify people, places, actions, or traits in...

...the input data and / or recorded data?

YES

NO

The technology is probably not a **surveillance tool** or an **automated decision system** - but plenty are!

The technology is probably an **automated decision system**, a type of **algorithmic system**.

The technology could be a **surveillance tool** but is probably not an **automated decision system**.

see AEKit's *Fill-in-the-Blank*



[ACLU-WA.org/AEKit/Fill-In](https://www.aclu-wa.org/AEKit/Fill-In)

see AEKit's *Questionnaire*



[ACLU-WA.org/AEKit/Questions](https://www.aclu-wa.org/AEKit/Questions)

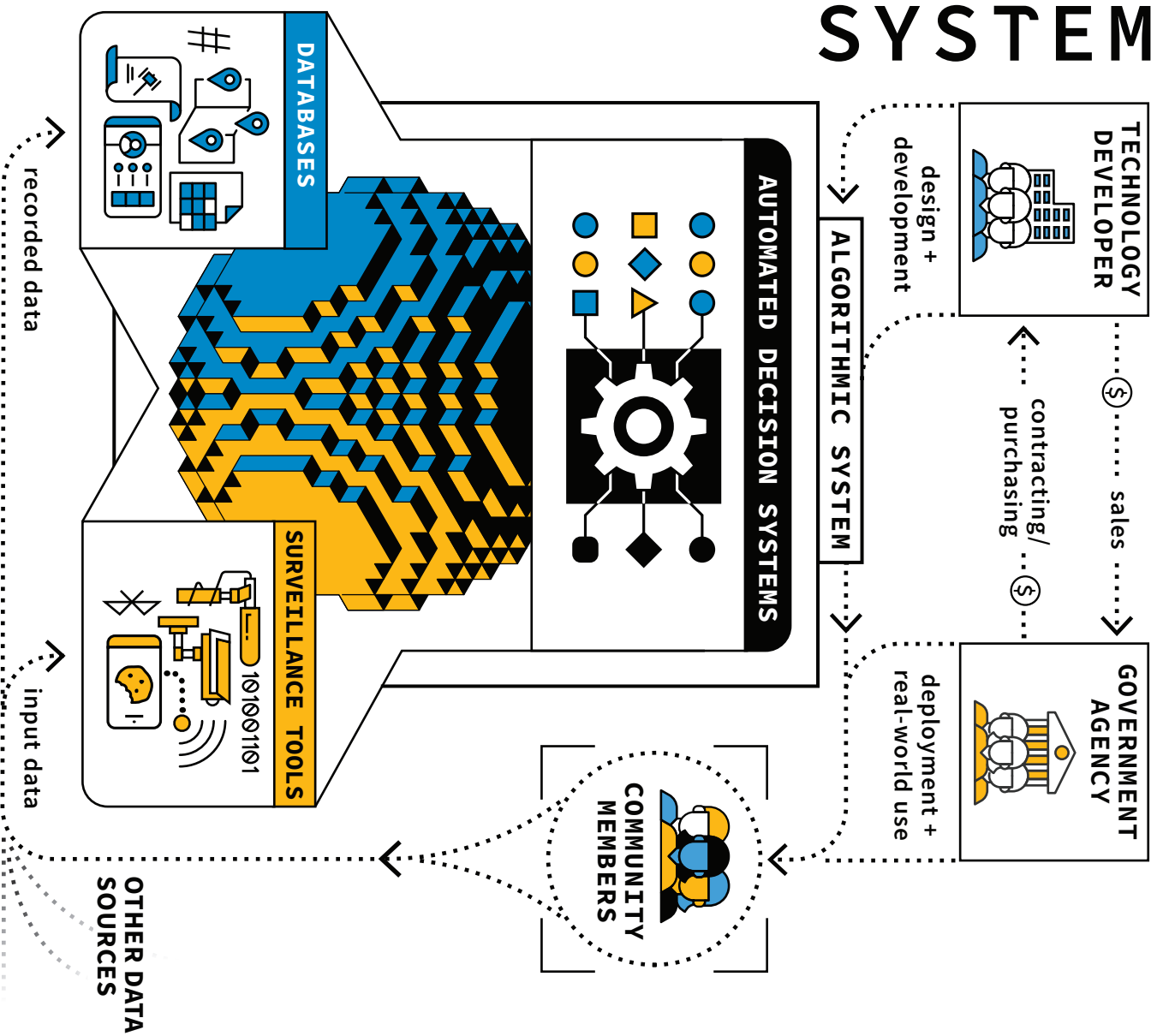
see ACLU's *They are Watching*



[www.TheyAreWatching.org](https://www.TheyAreWatching.org)

# MAP

# SYSTEM



# DEFINITIONS

## Input Data:

Words, photos, sounds, videos, clicks, gestures, location data or other signals or commands that a technology responds to or makes a record of.

## Surveillance Tools:

Technologies that collect, monitor, or curate information about people, behavior, or events in order to govern their behavior.

## Recorded Data:

A piece of information stored in or accessed by a technology. **Input data** becomes **recorded data** if it is stored by a technology.

## Databases:

Organized collections of **recorded data**.

## Algorithm:

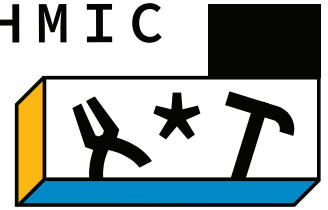
A precise and detailed set of steps or instructions that a computer follows to accomplish a task or function.

## Automated Decision System:

Any computerized system or process that aims to automate, aid, or replace human decision-making.

# FILL-IN-THE-BLANK

# ALGORITHMIC EQUITY TOOLKIT



## TO USE:

With a specific automated decision system in mind, fill in the blanks below with your best guesses and any information that is easily accessible online. Try to write a couple guesses for each of the larger blanks!

*Automated decision systems consist of multiple interrelated parts. The organizations responsible for these systems are often not transparent about this complexity. To intervene in the use of these systems, it can be useful to analyze different parts' origins, intents, and impacts.*

### NAMES

> I call the system ..... (short name)

> But it's officially called ..... (trade name)

### GOVERNMENT AGENCY

> The system is (or may be) used / deployed by: ..... (agency name)

Does this government agency have a report, briefing, or other documentation about the system?

- yes! I found it here: ..... (source)
- no...

### TECHNOLOGY DEVELOPER

> The system is provided / developed by: ..... (developer name)

Does this technology developer have something online that provides any information about the system?

- yes! I found it here: ..... (source)
- no...

### INPUT DATA

> The system detects, collects, and/or makes a record of:  
:  
&  
:  
..... (input data)

*For example: photos, sounds, videos, clicks, and / or location traces*



**RECORDED DATA**

> The system uses / references:

⋮  
&  
⋮

(recorded data)

For example: databases, maps, government statistics, laws + ordinances, and / or social media profiles

**INTENT**

> From the perspective of a software engineer at **TECHNOLOGY DEVELOPER**, the system does:

⋮  
&  
⋮

(technological function)

> I think that **GOVERNMENT AGENCY** may use the system to:

⋮  
&  
⋮

(intended use)

**POTENTIAL FOR ABUSE**

> I think the system could personally affect *me* if / when it is used for:

(personal impact)

> I think the system could affect:

.....  
if / when it is used for:

(a community that I care about)

(community impact)



What blanks were hard to fill? What information hasn't been properly disclosed? In an ideal world, how would you be able to access this information?



For more info, such as a flowchart to identify automated decision systems, scan this QR code with your phone camera to see the rest of the AEKit: [www.ACLU-WA.org/AEKit](http://www.ACLU-WA.org/AEKit)



# QUESTIONNAIRE

# ALGORITHMIC EQUITY TOOLKIT



## TO USE:

Ask these questions about automated decision systems (ADS) to government employees, elected officials, and vendors.



For more info, such as how to identify ADS, scan the QR code with your phone camera to see the rest of the AEKit. ([www.ACLU-WA.org/AEKit](http://www.ACLU-WA.org/AEKit))

*Automated decision systems make mistakes, and the types of mistakes they make can put people with marginalized identities at increased risk. These technologies are not always necessary. Some systems may be too invasive or risky by design, which is enough reason to reject a system outright.*

A).

## ACCURACY & ERROR IN ALGORITHMIC SYSTEMS

## GOALS:

**Policy makers should be able to demonstrate that:**

- The system won't make false or misleading assessments.
- People using the system are trained to recognize situations where false results are likely.
- Robust, auditable oversight of the system is in place.

Some technologies used by governments are inaccurate. They don't measure or detect what they claim to, or they do it poorly. This can result in decisions that adversely affect some individuals more than others. A single error in some contexts can result in a fatal or life-altering situation for a person from a historically marginalized community.

**EXAMPLE**—> *Automated license plate readers can misidentify letters and numbers on the plate or the state where the plate was issued.*



**EXAMPLE**—> *Facial recognition systems are never completely accurate; photos of suspects can be incorrectly matched with mugshots in a police database and falsely identify a person.*



A1:

**How accurate is the system? How often and under what conditions does it make mistakes? Does it have settings to adjust for more precise predictions?**

- > What evidence is there that the accuracy of the system has been independently tested, besides the manufacturer's claims?
- > How will the system perform in the local context where it is being deployed? Systems should be checked for their real-world performance in the places they are used.
- > How does the system perform when presented with diverse characteristics such as skin tone, lighting, signal interference, movement, or incomplete information?

## A2:

### What policies and procedures are in place when the system makes a mistake?

- > How are users of the system trained to recognize and resolve errors?
- > How do reporting processes publicly disclose errors when they occur?
- > What mechanisms are in place for auditing outcomes?
- > What is the role of community oversight in monitoring errors and outcomes?
- > What penalties exist for harms resulting from inaccurate assessments?
- > What protections are there for whistleblowers?

## B).

# INJUSTICE IN ALGORITHMIC SYSTEMS

Even when a system works perfectly accurately, it can still cause harm. The records that the system relies on can reflect previous discrimination, or the system can be applied in unjust ways.

**EXAMPLE** -> *Applicant tracking systems can replicate discriminatory hiring practices because of reliance on records of previous hiring.*



**EXAMPLE** -> *A 100% accurate facial recognition system could be used for harmful applications, such as identifying protestors.*



## GOAL:

Policy makers should be to explain how:

The system will not replicate historical patterns of bias like racism or sexism.

## B1:

### Where does the data that the system is using come from? Who gathered that data, with what tools, and for what purposes?

- > How has the data been audited to ensure it does not reflect discriminatory practices like racial profiling?
- > Will the data be re-purposed from the original reason it was collected? If so, how?

## B2:

### If the system works without errors, does it still perpetuate injustice?

- > What say do community members have in how the system is implemented (including where and when the system is used)? Can community members object and have their objections heard?
- > How can the public access and correct system records?
- > What are the explicitly intended and allowable uses of the system?
- > Are there oversight mechanisms in place to ensure the system is only being used for the specific purposes claimed? If so, what are they?
- > Are there any disciplinary penalties for misuse of the system? If so, what are they?