Red flags and checklist: Outline

Part 1: (what questions and why)

Appropriate use– the tool's ability to accurately predict recidivism, crime, etc.– i.e. a higher rate of predictive accuracy.

- Is there information the tool will be used only for the intended purpose and not to track individuals, for instance, will the Automated License Plate Readers (ALPR) be used to determine patterns and target drivers who visit gun shops, immigration clinics, health centers, protests or places of religious worship?
  Yes_____ No_____ N/A_____

- appropriate and compatible for the purpose intended, for instance, ….?
- Does the technology collect and stores your personal information?

Training Bias and Fairness/equity– the tool's ability to treat all users/communities fairly without racial bias, or discrimination the underlying influence of the data and the people who built the algorithm

[§ Is the data representative? does it include equal sample population from diverse voices/communities and context?
§ Is the data local and representative of the community?
§ Does the algorithm make final assessments and recommendations without an interference by a person?
§ Which data and variables were used in the algorithms?
§ Are variables such as race, gender, zip codes, age that are proxies for discrimination included in the algorithm?]

Impact – examine the algorithm in terms of effect it will have on the community members/users

§ What is the overall impact of the technology?

- Will the technology have negative impacts on the users such as denial of benefits or services?
  Yes_____ No_____ N/A_____

- Will the technology violate privacy for users by .....?

- Will the technology have an impact on minorities such as discrimination, ableism, the predpol replicates bias ? (reinforcing existing bias)
  Yes_____ No _____ N/A_____

- Will the technology provide positive changes to all community members?
  Yes_____ No_____ N/A_____

- Is the technology reinforcing or amplifying existing bias by targeting minority and low income communities, for instance leading to feedback loops?

  [§ What is the type of impact? i.e. access to benefits or services, financial, privacy, freedom, rights.
  § Who will be impacted by the technology?
  § What are the potential negative impacts of the technology on community/ community members based on age, race, religion, national origin, gender, disability?
  § Will the data follow users throughout their lifetime, affect their reputations and impact future opportunities?]

Transparency – the extent to which the algorithms (codes, data) used are available to users.

- Is there any information about the data used to create the technology, such as sample population, when it was collected, how it was analyzed, any limitations, uncertainty, accuracy?
  Yes _____ No_____ N/A_____

- Is there any information about the models and algorithms used for the technology?
  Yes_____ No _____ N/A_____

- Is the data used validated and representative of the community, real-world situations, for instance, if the technology is a pretrial risk assessment tool, was the training data representative of your community/jurisdiction?
  Yes_____ No _____ N/A_____

  [§ Is there any information about the data used in terms of accuracy, uncertainty, limitations, assumptions and the sample population?
  § How was the data collected, transformed and analyzed?

§ Which tools were used to model the data? What features or variables are used in the algorithm?
Does the technology reveal details about accuracy rates, margin of error and measures of uncertainty?]

Accountability –

§ Who or what made the decisions about the data used?
§ How were the decisions made?
§ Can you review or audit those decisions?
§ How was the algorithm tested before being used?
§ How is it tracked, or modified?

Interpretability – the extent to which the tool can be understood by users, government agencies, officials, stakeholders and community organizations

- Does the technology provide clear documentation on how to interpret the models and outputs?
  Yes_____ No _____ N/A_____
- Before making a final hiring decision using a tool, do you understand the data processing that led to the information being used?
  Yes_____No_____ N/A_____
- Are there policies or guidelines for proper use of the technology?
  Yes_____ No_____ N/A_____

Operability – the extent to which the tool can be administered by government agencies and officials. (implementation practices)

- Have you been trained how to operate the technology correctly?
  Yes_____ No_____ N/A_____
- Is there a straightforward and non-technical terms that describe the technology, its use, inputs and outcomes?
  Yes_____ No_____ N/A_____

[§ Will the agencies and/or organizations operating the technologies be trained how to use them?
§ Who will have access to the technology?]

Security & privacy –

- Is there information about the measures taken to protect personally identifiable data such as name, address, or face so that it does not lead to tracking the user or violating their privacy?
  Yes_____ No_____ N/A_____

- Will the data collected be used only for the purpose intended and not be shared with other agencies, government or companies? For instance, will the data collected by ALPR also be shared with banks, auto recovery companies, or insurance companies?
  Yes_____ No_____ N/A_____

- Is there information about data security, for instance how information shared amongst law enforcement agencies from the automated license plate readers will be protected?
  Yes_____ No_____ N/A_____

  § How will personal information be protected from data breaches?
  § Does the technology provide comparable privacy protections to different groups?
  § Was the data used private or public, are there any identifying aspects that could cause privacy harm to the individuals.?

<u>Part 2:</u>
**Add 2 -3 specific examples of technologies and their harms.

<u>Part 3 additional features/resources:</u>
AI Now. 2018. Algorithm accountability policy toolkit.
Bavitz, Christopher, Sam Bookman, Jonathan Eubank, Kira Hessekiel, and Vivek Krishnamurthy. 2018. Assessing the Assessments: Lessons from Early State Experiences in the
Procurement and Implementation of Risk Assessment Tools. *Berkman Klein Center for Internet & Society research publication.*
Center for Government Excellence. Ethics and Algorithms Toolkit https://ethicstoolkit.ai/

Diakopoulos N. 2016. Accountability in Algorithmic Decision Making. *Communications of the ACM, Vol. 59(2).*
https://cacm.acm.org/magazines/2016/2/197421-accountability-in-algorithmic-decision-making/fulltext

https://www.eff.org/pages/automated-license-plate-readers-alpr

https://www.rand.org/pubs/research_reports/RR2708.html

Algorithmic Equity Checklist

Possible harms	Yes	No	N/A
Impact			
The effect technology will have on the community members/users			
1. Will the technology have negative impacts on the users such as denial of benefits or services?			