All comments will be made public as-is, with no edits or redactions. Please be careful to not include confidential business or personal information, otherwise sensitive or protected information, or any information you do not wish to be posted.

Comment Template for First Public Draft of NIST SP 1270 "A Proposal for Identifying and Managing Bias in Artificial Intelligence"

Submit comments by September 10, 2021 to:
ai-bias@list.nist.gov

Comment #	Commenter organization	Commenter name	Paper Line # (if applicable)	Paper Section (if applicable)	Comment (Include rationale for comment)	Suggested change
1	National Fair Housing Alliance	Michael Akinwumi	197 - 200	section 1: introduction	The technical characteristics needed to cultivate trust in AI systems should include fairness, responsibility, and auditability. These principles are core to advancing ethics of AI or machine learning and they should be promoted as features required to cultivate trusts in AI systems. Please check: (1) Müller, Vincent C., "Ethics of Artificial Intelligence and Robotics", The Stanford Encyclopedia of Philosophy (Summer 2021 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2021/entries/ethics-ai/>. (2) https://www.fatml.org/resources/principles-for-accountable-algorithms	"... Working with the AI community, NIST has identified the following technical characteristics needed to cultivate trust in AI systems: accuracy, explainability and interpretability, privacy, reliability, robustness, safety, security, fairness, responsibility, and auditability - and that harmful biases are mitigated." We recommend that disparate impact testing and less discriminatory alternative (LDA) search be included under fairness.
2	National Fair Housing Alliance	Michael Akinwumi	219	section 1: introduction	Authors wrote that "not all types of bias are negative" without an example or a clarification of when a bias does not become adversarial to consumers.	Authors should add at least an example of when a bias does not constitute negative experience for consumers and also clarify types of bias that are not negative.
3	National Fair Housing Alliance	Michael Akinwumi	220	section 1: introduction	The report should focus on both harmful societal outcomes and harmful outcomes for individuals. The scale (either society or a small group of individuals such as communities of color) should not dictate what the focus should be when trust, bias or discrimination of AI systems are at issue.	"... focuses on biases present in AI systems that can lead to harmful outcomes for the society and individual consumers."
4	National Fair Housing Alliance	Michael Akinwumi	335 - 342	section 2: the challenge posed by bias in AI systems	Every AI or machine learning model has a lifespan. An AI system that claims to be responsible should have a mechanism that monitors the usefulness of an AI solution post-deployment.	Add a sentence that accounts for monitoring AI solutions post-deployment as a reason for "potential public distrust of AI related to bias in systems."

5	National Fair Housing Alliance	Michael Akinwumi		344	section 2: the challenge posed by bias in AI systems	Authors mentioned "in-place and in-development AI technologies and systems". It is not clear what the qualifiers mean.	Clarify the contextual difference, either with examples, between "in-place" and "in-development" AI technologies and systems.
6	National Fair Housing Alliance	Michael Akinwumi		415	section 4: identifying and managing bias in artificial intelligence	Figure 1 suggests that the cycle of AI system management ends at deployment. However, deploying an AI solution generally leads to changes in underlying data (data that the AI system acts on post-deployment may differ in scope and patterns from those used to develop the system). Hence, a component that monitors the behavior and performance of an AI solution post-deployment should be included in the proposed approach for managin AI bias.	Expand scope of the approach for managing AI bias to include Monitoring (post-deployment). The performance post-deployment could be used to make a decision to update the model.
7	National Fair Housing Alliance	Michael Akinwumi		424	section 4: identifying and managing bias in artificial intelligence	Data understanding and decisions about what model or set of models to train are usually decided in the pre-design stage. These concepts are currently missing in the section.	Devote a section to Data Understanding, Data Collection, and Model Choice (for example mis-specifying a regression problem as a classification problem by discretizing the continuous target variable could lead to bias in model outcomes)
11	National Fair Housing Alliance	Michael Akinwumi		510	section 4: identifying and managing bias in artificial intelligence	The design and development stage is extremely light on the role that algorithms or models play in driving bias and discrimination downstream the AI solution pipeline.	Authors should emphasize roles that key decisions during model development phase play in amplify bias and discrimination in AI solution pipeline: target or label creation; feature selection and feature transformation; model selection and hyperparameter tuning; model fitness - overfitting and underfitting; performance evaluation.
8	National Fair Housing Alliance	Michael Akinwumi		523	section 4: identifying and managing bias in artificial intelligence	Authors recommend "... taking context into consideration ..." under Design and Development Stage. However, contexts are usually set at the problem formulation stage of the pre-design phase. In addition, the referenced context only refers to data context.	Clarify the context in "optimization over context" or rather be more specific about the context being referred to. If authors meant business or social context that prioritizes fairness over accuracy, as suggested on lines 521 and 522, then an example more relevant that the ecological fallacy's should be cited.
9	National Fair Housing Alliance	Michael Akinwumi		562	section 4: identifying and managing bias in artificial intelligence	Similar to other sections of the report, the real-world example highlights proxy issues though it seems the intent is to call attentions to how decisions made at the model development stage may lead to bias and discrimination downstream.	Use a more suitable example to help practitioners appreciate how subjective decisions made during model development stage may exacerbate AI-driven bias and discrimination. In addition, authors should underscore needs for less discrimination alternatives that prioritize model fairness over accuracy during model development stage.
10	National Fair Housing Alliance	Michael Akinwumi		583	section 4: identifying and managing bias in artificial intelligence	In "Since many AI-based tools can skip deployment to a specified expert end user", it is not clear if authors meant users are the ones who deploy AI or that the deployment is often outsourced.	If authors meant to highlight bias risks of outsourcing model deployment, the sentence should clarify that.

12	National Fair Housing Alliance	Michael Akinwumi	636 - 640	section 4: identifying and managing bias in artificial intelligence	Arguments that compare pre-deployment and post-deployment model performances fit well under model monitoring.	Expand scope of the approach for managing AI bias to include Monitoring and move these lines to that section.
13	National Fair Housing Alliance	Michael Akinwumi			The report is framed as an approach that could be used for identifying and mitigating bias in AI solutions. As is, it lacks actionable steps or guidance for mitigating or identifying any bias in AI.	The utility of the report would be increased if there are guidelines to follow for identifying, analyzing, and mitigating bias in AI. The current report largely focuses on data representativeness with little to no content on how algorithms may lead to bias even in the unlikely situation of perfect or representative data. In addition, while authors do a great job of identifying different things that could go wrong with a model in production, many details are missing on how machine learning operations (MLOps) may cause unintended bias or discrimination post-deployment. For example, CI/CD (contnuous integration/continuous deployment) is a core component of MLOps and a wrong integration of it into the system that operationalizes an AI solution may impede algorithmic fairness downstream. In addition, every model has a lifespan and monitoring mechanism should be integral to any approach proposed for identifying bias or discrimination in an AI system.