A Participatory Learning Framework for Multi-Stakeholder Decision-Making

Vittoria Vineis[a,*], **Giuseppe Perelli**[a] **and Gabriele Tolomei**[a]

[a]Sapienza University of Rome
ORCID (Vittoria Vineis): https://orcid.org/0009-0003-6074-8344, ORCID (Giuseppe Perelli): https://orcid.org/0000-0002-8687-6323, ORCID (Gabriele Tolomei): https://orcid.org/0000-0001-7471-6659

Abstract. Conventional automated decision-support systems, often based on supervised learning, focus on predicting outcomes to recommend actions. However, they typically overlook the complexity of multi-actor environments, where diverse and conflicting stakeholder preferences must be balanced. At the same time, participatory AI approaches remain largely context-specific, limiting their broader applicability. To address these gaps, we propose a participatory framework that reframes decision-making as a multi-stakeholder optimization problem, using context-dependent reward functions to represent each actor's preferences. Our modular, model-agnostic framework employs k-fold cross-validation to fine-tune user-provided prediction models and evaluate decision strategies, including compromise functions that mediate stakeholder trade-offs. A synthetic scoring mechanism aggregates user-defined preferences across multiple metrics to rank strategies and select an optimal decision-maker for generating actionable recommendations on new data. Validated on two high-stakes real-world case studies, the framework consistently produces stakeholder-aware decisions that outperform purely predictive baselines across multiple metrics, while enhancing the transparency and accountability of AI-supported decision-making.

1 Introduction

Advances in artificial intelligence (AI) and machine learning (ML) have boosted the development of *automated decision-making (ADM) systems* in high-stakes domains such as healthcare, finance, and public policy [10]. However, the predominant *top-down* design paradigm and the traditional optimization-oriented setup for these systems are increasingly under scrutiny, particularly due to their tendency to overlook critical stakeholder perspectives and broader societal impacts [18]. In many cases, algorithmic solutions emphasize predictive accuracy above all else, neglecting or underperforming in aspects such as fairness [37], transparency [20], and the consideration of heterogeneous and potentially conflicting interests [28]. This narrow focus often leads to *impact-blind* recommendations that risk perpetuating historical biases and exacerbating inequities [4], highlighting the urgent need for more trustworthy, accountable systems [30, 16, 11]. For example, in healthcare, predictive models for diagnosis or resource allocation may fail to account for disparities in service access or patient demographics, disproportionately harming marginalized communities [35]. In finance, credit scoring models often embed biases from historical data, excluding individuals from economic opportunities [21] and ignoring broader socioeconomic effects.

Recent efforts to mitigate these issues often involve post-hoc fairness adjustments or the adoption of generalized fairness metrics (e.g., [33, 36]). While these approaches represent steps toward more equitable AI, they often fall short of addressing the nuanced, context-dependent priorities of real-world stakeholders. In particular, many current solutions remain *single-actor* in focus, thereby missing the opportunity to model and reconcile multiple, potentially conflicting objectives in a transparent manner. In this context, recent developments in *participatory approaches to AI design* offer promising avenues, but often lack versatility across use cases.

To address these gaps, we propose a novel, modular, and flexible multi-stakeholder decision-making framework that reinterprets both the offline training and online deployment phases of standard predictive model–based ADM systems as a *multi-actor participatory learning and decision-making process*. At its core, the framework is domain- and model-agnostic, integrating multiple actor perspectives through dynamic reward modeling and principled decision strategies. It ranks strategies across diverse evaluation metrics, explicitly balancing heterogeneous stakeholder preferences. Rather than replacing traditional prediction-based systems, it complements them by adding context-awareness, transparency, and a richer representation of alternatives—thereby enhancing the expressivity, interpretability, and ultimately, the accountability of decision-making processes. To the best of our knowledge, no prior work addresses this problem from a comparable perspective.

Our main contributions are as follows:
(i) We bridge foundational contributions from reward-based learning, game theory, welfare economics, computational social choice, and optimization *to advance the formalization of participatory AI solutions*.
(ii) We introduce a *theoretically grounded framework* that overcomes the limitations of traditional single-perspective[1] systems by systematically modeling and reconciling diverse, potentially conflicting objectives in context-aware ways.
(iii) We demonstrate the *effectiveness and generalizability of the proposed framework* through rigorous experiments on two real-world

* Corresponding Author. Email: vineis@diag.uniroma1.it

[1] In this work, we use "single-perspective" to describe systems that are not multi-stakeholder aware. We treat "stakeholder" and "actor" largely interchangeably, the former rooted in social science, the latter in formal modeling terminology.

case studies, showing how incorporating stakeholder diversity into the AI training pipeline improves decision outcomes compared to purely predictive baselines across multiple evaluation metrics.
(iv) To promote transparency and facilitate broader adoption, we provide *complete access to the source code and experimental setup*, enabling full reproducibility and adaptation to new use cases.[2] The remainder of the paper is organized as follows. Section 2 summarizes the relevant work. Section 3 introduces the proposed framework, analysing computational complexity and detailing scope and deployment considerations. Section 4 presents experiments on real-world high-stakes scenarios, while Section 5 concludes with key takeaways and outlines future research directions.

2 Related Work

The topics of *fairness and accountability* in ADM systems have garnered significant attention among scholars and practitioners due to their increasingly pervasive deployment in multiple domains [10] and their potential social impact [2]. Research in this area has extensively explored pre-, in-, and post-processing techniques to mitigate biases in data and modeling pipelines (for comprehensive reviews see, for instance, [33, 36]), contributing to the promotion of fairness in algorithmically supported decision systems. However, studies have demonstrated that fair algorithms alone cannot guarantee fairness in practice [19, 27], and aspects such as interpretability and fairness are inherently interdependent factors adding complexity to their operationalization in real-world AI systems [13, 40, 39, 26]. The multifaceted nature of fairness is further influenced by cultural and social contexts, complicating efforts to develop fairness frameworks that extend beyond pre-defined universal metrics [41]. Moreover, some authors argue that fairness-aware methods, while often reducing output biases, fail to address systemic inequities, highlighting the need for approaches that incorporate diverse stakeholder interests and broader societal impacts to promote more equitable outcomes [17]. Especially in dynamic and interactive decision-making settings, where multiple actors' interests come into play, *multi-agent systems* may offer a promising framework. Contributions to multi-agent reinforcement learning [48] and multi-agent multi-armed bandit frameworks [25] leverage, for instance, welfare functions to foster fairness in multi-agent decision settings. In this context, Wen et al. [45] advance these efforts by integrating feedback effects into Markov decision processes, enabling the modeling of dynamic, long-term impacts of decisions on fairness outcomes, as demonstrated through a loan approval scenario. Despite their valuable contribution to the integration of the presence of multiple actors in ADM systems, though, most approaches rely on predefined fairness definitions and require specific problem structures, limiting their adaptability to evolving stakeholder preferences and use-case-specific requirements.

Against this backdrop, *Participatory AI* has emerged as a significant paradigm for integrating diverse stakeholder perspectives throughout the AI lifecycle, offering opportunities to foster context-dependent fairness and promote accountability [6]. This paradigm emphasizes collaboration and co-creation, promoting inclusivity across both technical and non-technical domains [24, 5]. Contributions in this area span diverse fields, including healthcare [14], judicial systems [3], civic engagement [1], philanthropy [29], and urban planning [38]. More technical applications include collective debiasing [8], collaborative debugging [34], ranking with partial preferences [7] and web-based tools for democratizing ML workflows [47]. Collectively, these contributions reflect what has been described as a "participatory turn" in AI design [12].

However, current challenges such as the technical complexity of AI systems, structural and social barriers to participation, and power asymmetries hinder broader adoption and expose some applications to the risk of what has been described as "participation washing" [43]. In this regard, Maas and Inglés [32] have recently pushed this critique further, arguing that participatory AI discourse must expand beyond surface-level inclusion mechanisms to address the deeper legal, institutional, and political-economic arrangements that sustain power asymmetries between AI developers and affected stakeholders: they advocate for a shift beyond participatory *design* toward participatory *systems*, embedding stakeholder power directly into the processes of ongoing monitoring, adaptation, and auditing, thus ensuring meaningful and enforceable stakeholder influence over AI system behavior. Similarly, Feffer et al. [15] emphasize that participatory ML approaches often fall short by limiting participation to static preference elicitation, without embedding those preferences into the operational mechanisms of decision-making and governance. Nevertheless, while it is true that participatory AI must be addressed from a multi-stakeholder perspective, it is equally true that, from a technical standpoint, the need for meaningful, context-aware participation currently constitutes a bottleneck to the scalability and potential reach of participatory AI [12], creating a vicious cycle in which limited application hampers the development of richer participatory mechanisms, which in turn further constrains their adoption and broader impact. This underscores the pressing need for flexible and modular frameworks capable of adapting effectively across diverse domains and use cases.

3 The Participatory Learning and Decision Framework

In traditional prediction-oriented decision-making systems the goal is to recommend an action to be taken based on a set of features and a predicted outcome. In contrast to these approaches, which rely solely on outcome predictions, we formulate the task of suggesting an action as a *multi-actor decision-making problem*, where each actor has individual preferences over possible actions and outcomes.

Definition 1 (Multi-Stakeholder Decision-Making Problem)**.** *Let $\mathcal{I} = \{i_1, i_2, \ldots, i_n\}$ be a set of stakeholders or actors, where $|\mathcal{I}| = n$. We define an actor as any real or symbolic entity that is influenced by the decisions suggested by the system and, therefore, holds a direct stake in its resulting outputs. The decision space consists of a set of possible actions $\mathcal{A} = \{a_1, a_2, \ldots, a_k\}$, where $|\mathcal{A}| = k$, and a set of feasible outcomes $\mathcal{O} = \{o_1, o_2, \ldots, o_m\}$, where $|\mathcal{O}| = m$. Given a context $\boldsymbol{x} \in \mathcal{X}$ representing exogenous conditions (e.g., applicant attributes in a lending use-case), the system seeks to recommend an action $a \in \mathcal{A}$ that balances the preferences of all actors over the predicted outcomes.*

The framework flexibly accommodates both discrete and continuous action spaces $\mathcal{A}$ and outcome spaces $\mathcal{O}$, covering both classification- and regression-oriented formulations. For continuous outcomes, direct integration or discretization can be employed as needed, yielding a unified treatment across problem types.

Assumption 1 (Discrete or Discretizable Action Space)**.** *While the framework is theoretically general and capable of handling continuous action spaces, the formal derivations and computational complexity analysis in this work assume that the action space $\mathcal{A}$ is either discrete or can be discretized into a finite representative set*

[2] Supplementary material is openly available in archival form [arXiv:2502.08542].

without significant loss of fidelity. This ensures that the system can efficiently evaluate and compare candidate actions across all components, maintaining computational tractability.[3] *This assumption aligns with real-world scenarios, where the set of feasible actions is typically finite or effectively constrained.*

All components of the framework and its end-to-end workflow, as well as its computational overhead, are discussed below.

3.1 Core Components of the Framework

Learning Actor Preferences

A central element of our framework is the assignment of *actor-based rewards* to each action–outcome pair, extending conventional prediction-focused methods to address the diverse preferences of multiple stakeholders. Inspired by reward-based learning approaches, this mechanism leverages feedback signals to guide decision-making in a manner consistent with stakeholder priorities.

We formally distinguish two levels of training data: (i) the *base dataset* $T = \{(\boldsymbol{x}, a, o)\}$, containing only historically observed context–action–outcome triplets; and (ii) the *augmented reward dataset* $T^+ = \{(\boldsymbol{x}, a, o, r_i)\}_{a\in\mathcal{A},o\in\mathcal{O},i\in\mathcal{I}}$, which enriches T with stakeholder-provided or synthesized[4] reward signals r_i, extending coverage over all (or sampled) feasible action–outcome combinations. This augmented dataset allows the construction of the actor-specific reward function and reward matrix.

Definition 2 (Actor-Specific Reward Function)**.** *Given the multi-stakeholder decision-making problem described in Definition 1, the actor-specific reward function for each $i \in \mathcal{I}$ is defined as a mapping $R_i : \mathcal{X} \times \mathcal{A} \times \mathcal{O} \to [0, 1]$, which assigns a normalized desirability score to each action–outcome pair (a, o) under context $\boldsymbol{x}$, reflecting the individual preferences of actor i.*[5]

From a computational standpoint, $R_i(\boldsymbol{x}, a, o)$ can be derived via static domain-specific rules or, more generally, approximated using a learned function $q_i : \mathcal{X} \times \mathcal{A} \times \mathcal{O} \to [0, 1]$, which is trained on T^+ and predicts the reward each actor i would assign to $(\boldsymbol{x}, a, o)$, subject to bounded approximation error.

Definition 3 (Actor Reward Matrix)**.** *Given context $\boldsymbol{x} \in \mathcal{X}$, the actor reward matrix is:*

$$\mathbf{R}_i(\boldsymbol{x}) \in [0, 1]^{|\mathcal{A}|\times|\mathcal{O}|},$$

where each entry (u, v) is defined as $R_i(\boldsymbol{x}, a_u, o_v)$ for $a_u \in \mathcal{A}$ and $o_v \in \mathcal{O}$.

Remark 1 (Partial Reward Signal Coverage and Generalization)**.** *While Assumption 1 ensures that the action space $\mathcal{A}$ is discrete or discretizable, this does not imply that reward signals r_i are available for all possible combinations of actions and outcomes. In practice, the joint space $|\mathcal{A}| \times |\mathcal{O}|$ may only be partially covered by stakeholder elicitation due to resource or feasibility constraints.*

For small discrete spaces, the augmented dataset T^+ may include exhaustive or near-exhaustive coverage, enabling direct construction of the full reward matrices $\mathbf{R}_i(\boldsymbol{x})$. For large discrete or continuous (discretized) spaces, however, T^+ is typically built as a representative sampled subset. *In these cases, the learned reward model $q_i(\boldsymbol{x}, a, o)$ generalizes beyond observed combinations, thanks to the standard generalization assumptions in machine learning, under which a sufficiently expressive model trained on a representative subset can reliably approximate unobserved inputs.*[6]

Assumption 2 (Non-Adversarial Stakeholder Preferences)**.** *We assume that the reward signals r_i provided for each actor $i \in \mathcal{I}$ reflect genuine, non-adversarial preferences over action–outcome combinations. Specifically, we do not model strategic or adversarial behaviors in the elicitation or provision of these rewards. This assumption allows the learned reward models q_i to serve as faithful approximations of actors' true priorities.*[7]

Importantly, Assumption 2 does not alter the structure or validity of the proposed framework, since in adversarial settings, robustness can be incorporated by computing first-order bounds on the aggregate scoring function introduced later in the paper, enabling principled adaptation to strategic behavior. However, we leave the development and evaluation of such robustness mechanisms to future work, as the present contribution focuses on the formal presentation of the core framework.

Outcome Prediction Model

Another key component of the framework, which serves as the cornerstone of traditional ADM systems, is the outcome prediction model. This model learns to predict outcomes based on historical data comprising context vectors and past actions.

Definition 4 (Outcome Prediction Model)**.** *Given the setup of Definition 1, an outcome prediction model is a mapping:*

$$f : \mathcal{X} \times \mathcal{A} \to \Delta(\mathcal{O}), \quad \textit{or} \quad f : \mathcal{X} \times \mathcal{A} \to \mathbb{R},$$

where $\Delta(\mathcal{O})$ denotes the probability simplex over the discrete outcome space $\mathcal{O}$, and $\mathbb{R}$ denotes real-valued predictions for continuous outcomes. As in conventional supervised learning, f is trained on the base dataset T comprising historically observed context–action–outcome triplets.

Actor-specific Expected Rewards

Building on the predicted outcomes for each context, the framework leverages the learned reward models q_i, trained as outlined above, to estimate the desirability of each context–action–outcome configuration. Accordingly, all expected reward computations are expressed directly in terms of q_i.[8]

Definition 5 (Expected Reward Operator)**.** *Given actor $i \in \mathcal{I}$, context $\boldsymbol{x} \in \mathcal{X}$, and action $a \in \mathcal{A}$, the expected reward under the predicted outcome distribution f is:*

$$\mathbb{E}[q_i(a \mid \boldsymbol{x})] = \begin{cases} \sum_{o\in\mathcal{O}} P(o \mid \boldsymbol{x}, a; f) \cdot q_i(\boldsymbol{x}, a, o), \\ \int_{\mathcal{O}} q_i(\boldsymbol{x}, a, o)\, p(o \mid \boldsymbol{x}, a; f)\, do, \\ q_i(\boldsymbol{x}, a, \hat{o}), \end{cases}$$

[3] We remark that large or continuous action spaces may be addressed using established approximation techniques or optimization methods, depending on the specific structure of the problem and application domain.

[4] See Section 3.3 for details.

[5] The reward functions R_i are assumed normalized to $[0, 1]$ without loss of generality, ensuring comparability and bounded aggregation across actors.

[6] For instance, see [44, 23].

[7] Notably, when the reward signals r_i are derived through authentic bottom-up stakeholder consultation, the learning of the reward models q_i assumes particular significance: it functions not merely as a technical approximation step but as a principled mechanism for embedding genuine stakeholder preferences directly into the system's operational logic.

[8] When the true actor preferences R_i are fully specified or observable, the framework can incorporate them directly in place of q_i.

where the first line applies if $\mathcal{O}$ is discrete, the second if $\mathcal{O}$ is continuous, and the third if f is deterministic (i.e., produces a single predicted outcome $\hat{o}$).

By marginalizing over the predicted outcomes for each action, we derive the actor's expected reward vector $\mathbb{E}[q_i(a \mid \boldsymbol{x})]$ by collecting the expectations across all $a \in \mathcal{A}$.

Definition 6 (Actor Expected Reward Vector)**.** *Given actor $i \in \mathcal{I}$ and context $\boldsymbol{x} \in \mathcal{X}$, the expected reward vector over all actions is:*

$$\mathbb{E}[q_i(\cdot \mid \boldsymbol{x})] = \left[\mathbb{E}[q_i(a_1 \mid \boldsymbol{x})], \ldots, \mathbb{E}[q_i(a_{|\mathcal{A}|} \mid \boldsymbol{x})]\right]^\top \in [0,1]^{|\mathcal{A}|}. \tag{1}$$

Decision Strategies

Once the vectors of actor-specific expected rewards are computed based on the predicted outcomes, a set of decision strategies can be applied to derive the action suggested by the system. It is worth remarking that a crucial aspect of our proposed framework is that it does not impose any predetermined decision strategy a priori; rather, it systematically identifies and selects the strategy that optimally balances the user-specified set of evaluation metrics for the given use case, while explicitly incorporating the interests of stakeholders as expressed through their (learnt) reward signals.

Definition 7 (Decision Strategy)**.** *Building on previous definitions, a decision strategy is a mapping:*

$$D : [0,1]^{|\mathcal{I}| \times |\mathcal{A}|} \times \mathcal{P} \to \mathcal{A},$$

where $[0,1]^{|\mathcal{I}| \times |\mathcal{A}|}$ is the matrix of actor-specific expected rewards across actions, $\mathcal{P}$ is a parameter space with hyperparameters tied to the decision strategy, and the output is a selected action $a^* \in \mathcal{A}$. The full set of decision strategies is denoted $\mathcal{D} = \mathcal{C} \cup \mathcal{B}$, where $\mathcal{C}$ is the set of compromise functions and $\mathcal{B}$ is the set of baseline strategies.

Compromise Functions. In multi-actor settings, selecting a single action requires aggregating potentially competing preferences. The framework is designed to flexibly integrate a set of *compromise functions* $\mathcal{C}$, where each $C_j \in \mathcal{C}$ encodes a specific collective decision-making principle and systematically translates heterogeneous actor preferences into a unified decision. Importantly, compromise functions can be drawn from a wide range of disciplines, including game theory, computational social choice, welfare economics, and multi-objective optimization, providing a rich repertoire of aggregation rules that reflect diverse normative goals. [9]

Definition 8 (Compromise Function)**.** *Given Definition 6, a compromise function $C_j \in \mathcal{C}$ is defined as:*

$$C_j\left(\{\mathbb{E}[\boldsymbol{q}(a \mid \boldsymbol{x})]\}_{a \in \mathcal{A}}, \boldsymbol{p}\right) = \arg\max_{a \in \mathcal{A}} \Phi_j\left(\mathbb{E}[\boldsymbol{q}(a \mid \boldsymbol{x})], \boldsymbol{p}\right),$$

where $\Phi_j : [0,1]^{|\mathcal{I}|} \times \mathcal{P} \to \mathbb{R}$ is a scalar scoring function assigning a numerical score to each action, and $\boldsymbol{p} \in \mathcal{P}$ contains auxiliary parameters (e.g., actor weights, disagreement and ideal points).

Baseline strategies. The framework also includes two representative *baseline strategies* $\mathcal{B} \subseteq \mathcal{D}$, which serve as reference points for comparison but do not explicitly balance multi-actor preferences: (i) the *Outcome Predictor Baseline* B_{pred}, which mirrors traditional ADM systems that optimize predicted outcomes alone and, with discrete outcomes, selects $a^* = \arg\max_{a \in \mathcal{A}} P(o^* \mid \boldsymbol{x}, a)$ where $o^* = \arg\max_{o \in \mathcal{O}} P(o \mid \boldsymbol{x}, a)$; and (ii) the *Individual Reward Maximization Baseline* $B^{(i)}_{\text{max}}$, which selects $a^* = \arg\max_{a \in \mathcal{A}} \mathbb{E}[q_i(a \mid \boldsymbol{x})]$.

[9] Due to space constraints, a non-exhaustive selection of representative compromise principles is presented directly in the experimental section.

Evaluation Metrics and Optimal Action Selection.

We can now formalize how the framework evaluates decision functions and selects the optimal one based on a set of user-defined performance metrics.

Evaluation Metrics Let $\mathcal{M} = \{M_1, M_2, \ldots, M_z\}$ be a set of evaluation metrics, each capturing a distinct objective such as predictive performance, fairness, or domain-specific criteria. Given a dataset of T context–action–outcome triplets $\{(\boldsymbol{x}_t, a, o)\}$, the raw performance of a decision function $D \in \mathcal{D}$ under metric M_h is computed as:

$$\mathrm{P}(D, M_h) = \frac{1}{T^\alpha} \sum_{t=1}^{T} \sum_{a \in \mathcal{A}} \sum_{o \in \mathcal{O}} \Theta(D, M_h \mid \boldsymbol{x}_t, a, o),$$

where $\Theta(\cdot)$ captures the metric-specific contribution on triplet $(\boldsymbol{x}_t, a, o)$, and $\alpha \in \{0, 1\}$ distinguishes between averaging metrics and summing metrics.

To ensure comparability across heterogeneous metrics, raw scores are normalized on a $[0, 1]$ scale:

$$\tilde{P}(D, M_h) = \frac{\mathrm{P}(D, M_h) - \min_{D'} \mathrm{P}(D', M_h)}{\max_{D'} \mathrm{P}(D', M_h) - \min_{D'} \mathrm{P}(D', M_h)},$$

where the minimum and maximum are taken over all $D' \in \mathcal{D}$. Note that depending on the semantic direction of each metric — whether higher values, lower values, or proximity to a defined target (e.g., minimal fairness gaps) are preferred — raw scores may be transformed prior to normalization.[10]

Aggregated Scoring and Optimal Strategy Selection Given a user-defined weight vector $\boldsymbol{w} = (w_1, \ldots, w_z)$ with $w_h \geq 0$ and $\sum_h w_h = 1$, the aggregated score is:

$$S(D) = \sum_{h=1}^{z} w_h \, \tilde{P}(D, M_h),$$

and the optimal decision function is selected as:

$$D^* = \arg\max_{D \in \mathcal{D}} S(D). \tag{2}$$

To ensure robust evaluation, the framework employs standard validation procedures (e.g., k-fold cross-validation), integrating them into the hyperparameter optimization process for the outcome prediction model and averaging metric performance across held-out splits.

Remark 2 (Optional Multi-Strategy Output)**.** *Although the framework identifies D^* as the top-ranked decision function, it can optionally present the full ranked list $\{D_{(1)}, D_{(2)}, \ldots, D_{(|\mathcal{D}|)}\}$, allowing users to explore alternative strategies and understand trade-offs across metrics.*

Online deployment. At inference time, given a new context $\boldsymbol{x'} \in \mathcal{X}$, the system computes actor-specific expected rewards $\{\mathbb{E}[q_i(a \mid \boldsymbol{x'})]\}_{i=1}^{N}$ across all $a \in \mathcal{A}$, and recommends:

$$a^* = D^*\left(\{\mathbb{E}[q_i(a \mid \boldsymbol{x'})]\}_{i=1}^{N}\right). \tag{3}$$

We summarize the core stages of the proposed participatory training and decision framework in Algorithm 1, detailing its offline learning and online deployment steps.

[10] Specifically, for gap-based or target-centered metrics, we compute the absolute deviation from the optimal target to ensure that, after normalization, higher normalized values uniformly correspond to better performance across all metrics.

Algorithm 1 Participatory Training and Decision Framework

Input: Base training dataset T, augmented reward dataset T^+; actor set $\mathcal{I}$; action space $\mathcal{A}$; outcome space $\mathcal{O}$; decision strategy set $\mathcal{D}$; evaluation metric set $\mathcal{M}$.
Output: Optimal decision strategy D^*; recommended action a^* for new context $\boldsymbol{x'}$.

```
[Offline Phase: Model Training and Strategy Selection]
Train outcome predictor f on T
for each actor i ∈ I do
    Train reward model q_i on T^+_sampled ⊆ T^+
end for
for each decision strategy D ∈ D do
    Compute aggregated performance score S(D) using metrics M over T
end for
Select optimal decision strategy D* (see Eq.2)
[Online Phase: Action Recommendation]
for each actor i ∈ I do
    Compute expected reward vector E[q_i(· | x')] (see Eq. 1)
end for
Select recommended action a* (see Eq. 3)
return a*
```

3.2 Computational Complexity Analysis

Theorem 1 (Additional Computational Overhead)**.** *Under Assumption 1, relative to a baseline outcome-prediction system, the additional computational overhead introduced by the participatory multi-actor frameworks, added on top of the base cost, is:*
(i) Offline (training and selection phase, per cross-validation run):

$$O\Big(|\mathcal{I}|\cdot\big(|G_q|\cdot c^q_{\text{train}}+T_{\text{val}}\cdot|\mathcal{A}|\cdot(c^q_{\text{inf}}+|\mathcal{D}|)\big)\;+\;T_{\text{val}}\cdot|\mathcal{D}|\cdot|\mathcal{M}|\Big)$$

where $|G_q|$ is the reward model hyperparameter grid size, T_{val} the validation set size, c^q_{train} the per-actor reward model training cost, and c^q_{inf} the per-actor inference cost.

(ii) Online (inference phase, per deployment context):

$$O\Big(|\mathcal{A}|\cdot|\mathcal{I}|\cdot(c^q_{\text{inf}}+1)\Big)\qquad (\textit{best } D^*),$$
$$O\Big(|\mathcal{A}|\cdot|\mathcal{I}|\cdot(c^q_{\text{inf}}+|\mathcal{D}|)\Big)\quad (\textit{all } D\in\mathcal{D}).$$

The dominant terms arise from actor-specific reward computations and multi-strategy evaluations. In real-world deployments, where the number of actors, actions, strategies, and metrics is typically moderate, the added overhead remains practically tractable.

For space constraints, the detailed proof is provided as supplementary material.

Remark 3 (Effect of Non-Convexity)**.** *While certain compromise objective functions Φ_j may involve non-convex optimization and require iterative or heuristic approximations in general settings, under Assumption 1, the worst-case per-context selection cost remains bounded by $O(|\mathcal{A}|)$ via exhaustive enumeration over candidate actions. As a result, non-convexity does not alter the asymptotic overhead expression derived above.*

3.3 Remark on Scope and Practical Deployment

Before presenting experimental results, we stress that the framework is deliberately general and modular, offering a structured basis for integrating stakeholder preferences into automated recommendation pipelines. While we specify how actor-specific reward models enter the decision process, the concrete instantiation of components—choice of function approximators, data augmentation strategies for training, methods of reward elicitation, evaluation metrics, or decision functions—remains open. This flexibility is intentional, ensuring adaptability across domains.

The effective use of the framework ultimately depends on the design of actor-specific reward signals, ideally rooted in authentic stakeholder input or objective impact measures. Addressing the broader challenge of aligning revealed and latent preferences is important but beyond our present scope. Our aim here is to study how participatory decision-making can be simulated at inference time, given access to stakeholder-informed reward functions. The framework is thus agnostic to the provenance of these signals, while remaining compatible with future advances in preference elicitation and modeling.

4 Experiments

Although the core contribution of this paper is the formal specification of the framework, we complement it with illustrative experiments to explore its behavior and applicability across real-world decision-making settings. In particular, we provide empirical evidence of the framework's potential in two representative high-stakes domains, demonstrating its flexibility, utility, and capacity to support stakeholder-aligned decisions. Additionally, we conduct ablation analyses to examine how variations in reward structure, training sample size, and predictive model capacity influence system performance across a range of evaluation metrics.

4.1 Real-World Use Cases

We present two use cases, namely a loan approval scenario (multi-classification) and a health treatment selection scenario (causal inference regression). In both experimental settings, reward structures are predefined using prototypical stakeholder heuristics with added uniform noise to simulate real-world variability and test model robustness. This experimental setup serves two main goals: assessing the framework's versatility across diverse domains and ensuring interpretability through a controlled environment, offering insight into its structural behavior. In both experimental scenarios, we evaluate an illustrative set of compromise-based decision strategies — namely *Maximin*, *Proportional Fairness*, *Kalai-Smorodinsky solution*, *Nash Bargaining solution*, and *Compromise Programming L2*— alongside single-actor strategies and predictive-only baselines.

Use Case 1: Lending Scenario

We apply the framework to real-world data from the publicly available Lending Club dataset, structured as a 3×3 problem where the decision space $\mathcal{A}$ includes three actions ("Grant," "Grant lower amount," "Not Grant") and the outcome space $\mathcal{O}$ consists of three repayment states ("Fully Repaid," "Partially Repaid," "Not Repaid"). Context features include applicant-specific attributes such as credit score, income, financial history, and demographics.

This setting models three key stakeholder groups:

- *Bank*: seeks to maximize profitability, with rewards tied to repayment outcomes.
- *Applicant*: prioritizes loan access, with rewards reflecting the utility of approval balanced against repayment obligations.

- *Regulatory Body*: aims to ensure financial stability and inclusivity, placing value on equitable credit access for vulnerable groups.

Beyond baseline strategies, we benchmark compromise functions against an *Oracle* strategy simulating an idealized bijective mapping between outcomes and optimal actions. Performance is evaluated using standard fairness and accuracy metrics, as well as domain-specific measures such as profit percentage, loss percentage, and unrealized profit due to suboptimal decisions—each measured in relative terms with respect to the Oracle benchmark. Both the outcome prediction and actor reward models are implemented as Random Forests[11], and for ablation, we include a simpler k-Nearest Neighbor baseline.

Use Case 2: Healthcare Scenario

For the healthcare scenario, we use the first realization of the Infant Health and Development Program (IHDP) dataset, as introduced by Hill [22] and widely employed in causal inference research [42, 31, 46]. This dataset originates from a randomized experiment studying the effect of home visits on infant cognitive test scores.

The decision space comprises two actions (treatment vs. no treatment), with outcomes given by continuous cognitive scores. The context vector includes 25 binary and continuous features describing child and family characteristics.

This scenario illustrates the framework's applicability to causal inference tasks, modeling three stakeholder groups:

- *Healthcare Provider*: aims to improve patient outcomes while managing costs.
- *Policy Maker*: seeks to maximize societal benefit and ensure equity across demographic groups.
- *Parent*: values direct improvements in child well-being.

In addition to individual maximization baselines, we include a strategy that maximizes overall cognitive scores which naturally recommends treatment for all potential patients. Case-specific evaluation metrics include the mean outcomes for treated versus control groups and their absolute difference. To estimate the Conditional Average Treatment Effect (CATE), we use an X-regressor meta-learner [9] with XGBRegressor as the base learner.

4.2 Discussion of Key Insights

Although the ultimate objective of the framework is to identify, offline, the most suitable decision strategy for inference-time deployment, we conduct a comparative assessment of multiple decision functions to illustrate how variations in design choices affect performance across diverse evaluation dimensions.

Figures 1 and 2 provide a comprehensive empirical breakdown of how variations in reward structures and outcome model complexity shape decision function performance across predictive, fairness, and case-specific metrics in the lending scenario. All results are averaged over four random seeds and benchmarked against an *Oracle* strategy to reveal upper-bound potential for outcome prediction. A key observation across all experiments is that baseline strategies tend to optimize classical predictive metrics, notably accuracy, but underperform in fairness dimensions such as *Demographic_Parity* and in the diversity of actions recommended. More broadly, the observed performance differences between compromise-based strategies and single-actor or single-objective baselines underscore the inherent limitations of narrowly optimized systems: while such baselines may perform well on isolated metrics, they consistently underperform when assessed across a broader range of evaluation dimensions, particularly those capturing equity and diversity. Critically, the magnitude of the performance gap between narrow-objective solutions and multi-actor compromise approaches reveals the limited capacity of traditional single-perspective systems to reconcile competing priorities — and, by extension, their diminished ability to deliver socially informed, context-sensitive, and accountable decisions.

Figure 1 reveals how shifting the underlying actor reward functions — from balanced to more self-interested formulations[12] — sharpens the performance–fairness trade-off. This underscores the subtle but powerful role that reward design plays in shaping aggregate system behavior.

Figure 2 and 3 together allow observing the impact of increasing predictive power — whether through greater model expressiveness or larger training sample size — on decision strategy performance. The figures show that while improvements in performance-oriented metrics for single-objective strategies are evident (see *Accuracy* and *Total_Profit*), they are not accompanied by proportionally large gains in other dimensions, nor do they substantially and unambiguously alter the relative performance of compromise-based strategies. This emphasizes, on one hand, the need to evaluate systems across a broader set of dimensions beyond predictive accuracy alone, and on the other hand, that when balancing diverse stakeholder preferences, simply optimizing the predictive power of the underlying model remains insufficient.

Finally, Figure 4 refers to the healthcare scenario and demonstrates how the proposed framework can be used also to evaluate the distribution of treatment effects across patients, based on treatment assignment decisions and the expected treatment effects. This analysis underscores the framework's ability to assess not only traditional metrics associated with causal effects but also the equity and impact of treatment allocation across different population groups.

Taken together, these insights highlight the versatility of the proposed framework across domains and its potential to deliver stakeholder-sensitive and context-aware explainable decisions — even when operating on identical predictive inputs and without relying on pre-imposed or rigid fairness constraints.

5 Conclusion and Future Work

Traditional data-driven systems typically recommend actions based solely on predictive accuracy or fixed performance metrics, which — even when constrained by specific fairness criteria — often overlook the diverse and sometimes conflicting preferences of stakeholders. In this work, we introduced a flexible and model-agnostic learning framework that integrates heterogeneous stakeholder interests into the automated decision-making process, effectively simulating a participatory process whenever new decisions must be made. By combining actor-specific reward modeling with compromise strategies, the framework enables the balancing of competing priorities while optimizing user-defined evaluation metrics across domains.

The experiments demonstrate how the proposed framework can serve as a flexible tool to enhance the context-awareness, account-

[11] We use Random Forests for their robustness with non-differentiable reward signals and their wide practical adoption.

[12] In *Strictest Strategy* setting in Figure 1, the *Bank* values loan approvals only when full repayment is expected, while the *Applicant* prioritizes loan approval regardless of repayment likelihood.

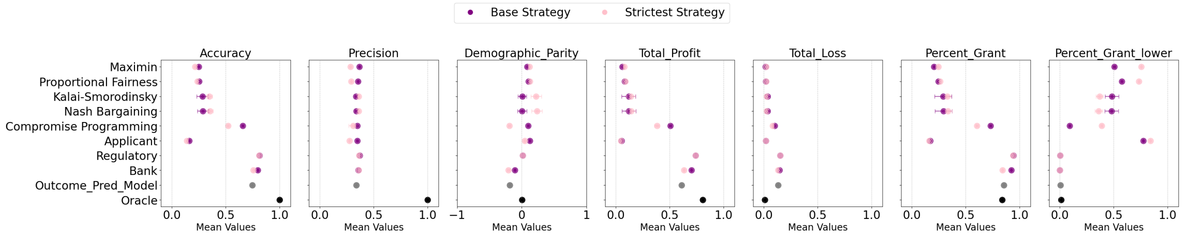

Figure 1. Comparison of decision strategy performance on the test set across evaluation metrics in the lending scenario, under varying reward structure

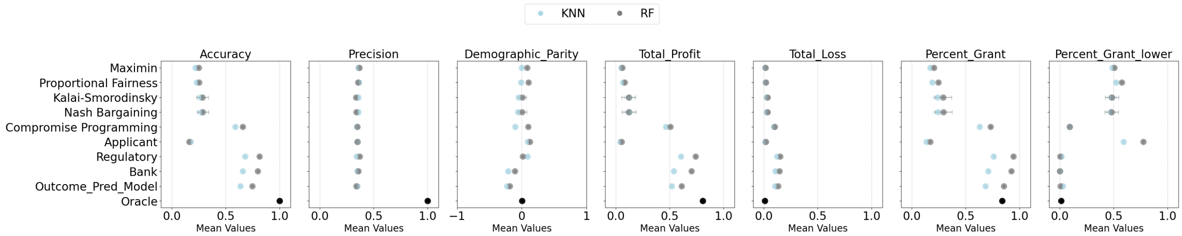

Figure 2. Comparison of decision strategy performance on the test set across evaluation metrics in the lending scenario, under varying predictive models

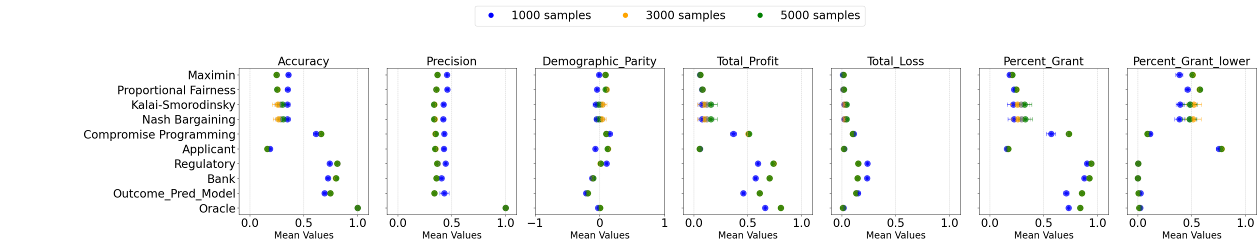

Figure 3. Comparison of decision strategy performance on the test set across evaluation metrics in the lending scenario, under varying training sample size

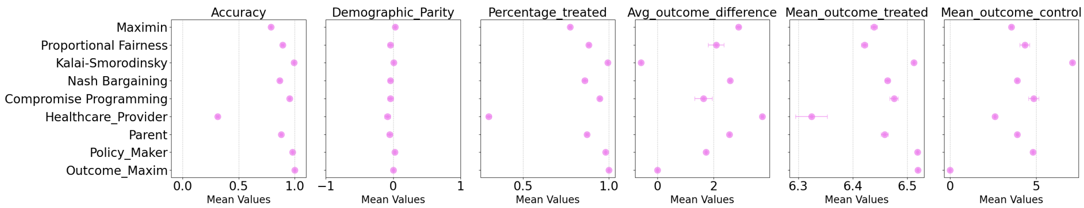

Figure 4. Comparison of decision strategy performance on the test set across evaluation metrics in the healthcare scenario

ability and transparency of automated decision-support systems, empowering practitioners to explicitly trace how stakeholder preferences shape final decisions and to justify trade-offs in high-stakes environments, without requiring rigid fairness constraints or modifications to the underlying predictors. As future work, we plan to relax Assumption 2, analyzing the framework's robustness and strategy-proofness in adversarial settings where stakeholders may strategically adjust their stated preferences to influence collective decisions. Similarly, we aim to study how biases, embedded both in the training data and within the stakeholder-provided reward signals, propagate through the system and affect final decisions. Promising directions also include extending the framework to accommodate evolving stakeholder preferences over time and deploying it in real-world contexts involving active preference elicitation and user evaluation.

Acknowledgements

This work was partially supported by: SERICS (PE00000014) under the National Recovery and Resilience Plan funded by the European Union—NextGenerationEU; HyperKG—Hybrid Prediction and Explanation with Knowledge Graphs (2022Y34XNM), funded by the Italian Ministry of University and Research (PRIN 2022); GHOST—Protecting User Privacy from Community Detection in Social Networks (B83C24007070005), funded by Sapienza University of Rome ("Progetti di Ricerca Grandi"); ASGARD—Autonomous and Self-Governing Agent-Based Rule Design (B83C23005800001), funded by Sapienza University of Rome ("Progetti di Ricerca Grandi").

References

[1] M. Arana-Catania, F. A. V. Lier, R. Procter, N. Tkachenko, Y. He, A. Zubiaga, and M. Liakata. Citizen participation and machine learning for a better democracy. *Digital Government: Research and Practice*, 2(3): 1–22, 2021.

[2] T. Araujo, N. Helberger, S. Kruikemeier, and C. H. De Vreese. In ai we trust? perceptions about automated decision-making by artificial intelligence. *AI & society*, 35(3):611–623, 2020.

[3] C. Barabas, C. Doyle, J. B. Rubinovitz, and K. Dinakar. Studying up: reorienting the study of algorithmic fairness around issues of power. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pages 167–176, 2020.

[4] S. Barocas and A. D. Selbst. Big data's disparate impact. *Calif. L. Rev.*, 104:671, 2016.

[5] A. Berditchevskaia, E. Malliaraki, and K. Peach. *Participatory AI for humanitarian innovation*. Nesta, London, 2021.

[6] A. Birhane, W. Isaac, V. Prabhakaran, M. Diaz, M. C. Elish, I. Gabriel, and S. Mohamed. Power to the people? opportunities and challenges for participatory ai. In *Proceedings of the 2nd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*, pages 1–8, 2022.

[7] K. Cachel and E. Rundensteiner. Prefair: Combining partial preferences for fair consensus decision-making. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency*, pages 1133–1149, 2024.

[8] E. Chan, Z. Liu, R. Qiu, Y. Zhang, R. Maciejewski, and H. Tong. Group fairness via group consensus. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency*, pages 1788–1808, 2024.

[9] H. Chen, T. Harinen, J.-Y. Lee, M. Yung, and Z. Zhao. Causalml: Python package for causal machine learning. *arXiv preprint arXiv:2002.11631*, 2020.

[10] F. Chiusi, B. Alfter, M. Ruckenstein, and T. Lehtiniemi. Automating society report 2020. 2020.

[11] E. Council. Regulation (eu) 2024/1689 of the european parliament and of the council of 13 june 2024 laying down harmonised rules on artificial intelligence and amending regulations (ec) no 300/2008,(eu) no 167/2013,(eu) no 168/2013,(eu) 2018/858,(eu) 2018/1139 and (eu) 2019/2144 and directives 2014/90/eu,(eu) 2016/797 and (eu) 2020/1828 (artificial intelligence act) off. *J. Eur. Union*, 50:202, 2024.

[12] F. Delgado, S. Yang, M. Madaio, and Q. Yang. The participatory turn in ai design: Theoretical foundations and the current state of practice. In *Proceedings of the 3rd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*, pages 1–23, 2023.

[13] J. Dodge, Q. V. Liao, Y. Zhang, R. K. Bellamy, and C. Dugan. Explaining models: an empirical study of how explanations impact fairness judgment. In *Proceedings of the 24th international conference on intelligent user interfaces*, pages 275–285, 2019.

[14] J. Donia and J. A. Shaw. Co-design and ethical artificial intelligence for health: An agenda for critical research and practice. *Big Data & Society*, 8(2):20539517211065248, 2021.

[15] M. Feffer, M. Skirpan, Z. Lipton, and H. Heidari. From preference elicitation to participatory ml: A critical survey & guidelines for future research. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, pages 38–48, 2023.

[16] L. Floridi et al. *Ethics, governance, and policies in artificial intelligence*. Springer, 2021.

[17] A. Gerdes. A participatory data-centric approach to ai ethics by design. *Applied Artificial Intelligence*, 36(1):2009222, 2022.

[18] F. Gerdon, R. L. Bach, C. Kern, and F. Kreuter. Social impacts of algorithmic decision-making: A research agenda for the social sciences. *Big Data & Society*, 9(1):20539517221089305, 2022.

[19] N. Goel, A. Amayuelas, A. Deshpande, and A. Sharma. The importance of modeling data missingness in algorithmic fairness: A causal perspective. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 7564–7573, 2021.

[20] S. Grimmelikhuijsen. Explaining why the computer says no: Algorithmic transparency affects the perceived trustworthiness of automated decision-making. *Public Administration Review*, 83(2):241–262, 2023.

[21] M. Hardt, E. Price, and N. Srebro. Equality of opportunity in supervised learning. *Advances in neural information processing systems*, 29, 2016.

[22] J. L. Hill. Bayesian nonparametric modeling for causal inference. *Journal of Computational and Graphical Statistics*, 20(1):217–240, 2011.

[23] K. Hornik, M. Stinchcombe, and H. White. Multilayer feedforward networks are universal approximators. *Neural networks*, 2(5):359–366, 1989.

[24] S. Hossain and S. I. Ahmed. Towards a new participatory approach for designing artificial intelligence and data-driven technologies. *arXiv preprint arXiv:2104.04072*, 2021.

[25] S. Hossain, E. Micha, and N. Shah. Fair algorithms for multi-agent multi-armed bandits. *Advances in Neural Information Processing Systems*, 34:24005–24017, 2021.

[26] A. Jain, M. Ravula, and J. Ghosh. Biased models have biased explanations. *arXiv preprint arXiv:2012.10986*, 2020.

[27] H. Jeong, H. Wang, and F. P. Calmon. Fairness without imputation: A decision tree approach for fair prediction with missing values. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 9558–9566, 2022.

[28] B. Laufer, T. Gilbert, and H. Nissenbaum. Optimization's neglected normative commitments. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, pages 50–63, 2023.

[29] M. K. Lee, D. Kusbit, A. Kahng, J. T. Kim, X. Yuan, A. Chan, and A. D. Procaccia. Webuildai: Participatory framework for algorithmic governance. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–35, 2019.

[30] B. Lepri, N. Oliver, E. Letouzé, A. Pentland, and P. Vinck. Fair, transparent, and accountable algorithmic decision-making processes: The premise, the proposed solutions, and the open challenges. *Philosophy & Technology*, 31:611–627, 2018.

[31] C. Louizos, U. Shalit, J. M. Mooij, D. Sontag, R. Zemel, and M. Welling. Causal effect inference with deep latent-variable models. *Advances in neural information processing systems*, 30, 2017.

[32] J. Maas and A. M. Inglés. Beyond participatory ai. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, volume 7, pages 932–942, 2024.

[33] N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan. A survey on bias and fairness in machine learning. *ACM computing surveys (CSUR)*, 54(6):1–35, 2021.

[34] Y. Nakao, S. Stumpf, S. Ahmed, A. Naseer, and L. Strappelli. Toward involving end-users in interactive human-in-the-loop ai fairness. *ACM Transactions on Interactive Intelligent Systems*, 12(3):1–30, 2022.

[35] Z. Obermeyer, B. Powers, C. Vogeli, and S. Mullainathan. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464):447–453, 2019.

[36] D. Pessach and E. Shmueli. A review on fairness in machine learning. *ACM Computing Surveys (CSUR)*, 55(3):1–44, 2022.

[37] E. Petersen, M. Ganz, S. Holm, and A. Feragen. On (assessing) the fairness of risk score models. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, pages 817–829, 2023.

[38] S. J. Quan, J. Park, A. Economou, and S. Lee. Artificial intelligence-aided design: Smart design for sustainable city development. *Environment and Planning B: Urban Analytics and City Science*, 46(8):1581–1599, 2019.

[39] R. Ramachandranpillai, R. Baeza-Yates, and F. Heintz. Fairxai-a taxonomy and framework for fairness and explainability synergy in machine learning. *Authorea Preprints*, 2023.

[40] J. Schoeffer, M. De-Arteaga, and N. Kuehl. On the relationship between explanations, fairness perceptions, and decisions. *arXiv preprint arXiv:2204.13156*, 2022.

[41] A. D. Selbst, D. Boyd, S. A. Friedler, S. Venkatasubramanian, and J. Vertesi. Fairness and abstraction in sociotechnical systems. In *Proceedings of the conference on fairness, accountability, and transparency*, pages 59–68, 2019.

[42] U. Shalit, F. D. Johansson, and D. Sontag. Estimating individual treatment effect: generalization bounds and algorithms. In *International conference on machine learning*, pages 3076–3085. PMLR, 2017.

[43] M. Sloane, E. Moss, O. Awomolo, and L. Forlano. Participation is not a design fix for machine learning. In *Equity and Access in Algorithms, Mechanisms, and Optimization*, pages 1–6, 2022.

[44] V. N. Vapnik. An overview of statistical learning theory. *IEEE transactions on neural networks*, 10(5):988–999, 1999.

[45] M. Wen, O. Bastani, and U. Topcu. Algorithms for fairness in sequential decision making. In *International Conference on Artificial Intelligence and Statistics*, pages 1144–1152. PMLR, 2021.

[46] L. Yao, S. Li, Y. Li, M. Huai, J. Gao, and A. Zhang. Representation learning for treatment effect estimation from observational data. *Advances in neural information processing systems*, 31, 2018.

[47] A. Zhang, O. Walker, K. Nguyen, J. Dai, A. Chen, and M. K. Lee. Deliberating with ai: improving decision-making for the future through participatory ai design and stakeholder deliberation. *Proceedings of the ACM on Human-Computer Interaction*, 7(CSCW1):1–32, 2023.

[48] M. Zimmer, C. Glanois, U. Siddique, and P. Weng. Learning fair policies in decentralized cooperative multi-agent reinforcement learning. In *International Conference on Machine Learning*, pages 12967–12978. PMLR, 2021.