Frameworks, Methods and Shared Tasks:

Connecting Participatory and Trustworthy AI Through a Systematic Review of Global Projects

Maria Isabel Magana Arango and Katie Shilton

As the discourse on responsible and trustworthy AI intensifies, Participatory AI (PAI) presents a compelling approach to the democratic development of automated technologies. But how should the field think about how and whether participatory methods increase trust in, and the trustworthiness of, AI systems? In response to the recent growth in PAI research, we conducted a systematic examination to understand the landscape of methods and theoretical lenses used in participatory AI projects, and to connect those methods and lenses to trust building. This paper analyzes 95 global PAI projects to understand the participatory landscape of trustworthy AI. Our review explores differences in theoretical frameworks, participation methods, and the details of *shared tasks* within the AI lifecycle across sectors and geographies. Our findings reveal an evolving definition of PAI, with actors implementing diverse methods and shared tasks. Focusing on shared tasks provides a lens for analyzing how participation can build trust in, and trustworthiness of, AI systems. Considering PAI through the lens of trust creates a framework for mapping PAI activities to both participant and broader public trust, and illustrates gaps where PAI processes alone will not be enough to ensure trustworthy AI systems.

Keywords: Artificial Intelligence, Participatory Artificial Intelligence, Participatory Methodologies, Trustworthy Artificial Intelligence

1 INTRODUCTION

Trust is a central problem for Artificial Intelligence (AI), and particularly for the generative models currently transforming the field. Because AI systems seek to augment or replace human decision-making with machine learning [52], because the stochasticity of generative models can be opaque and difficult to explain [15], and because the data used to train machine learning models is frequently rife with human biases [3] or an imperfect indicator for unobservable constructs [30], today's AI systems are frequently untrusted by experts or users [1]. Building trust is a key issue for AI design, adoption and governance, and is a key focus of a wide variety of national and international guidance for designing AI [43, 44].

Resonant with, but not fully intersecting with, discussions of trust in AI has beena surge in support for *participatory AI* (PAI) approaches. These methods intend to incorporate the public into designing, developing, evaluating, and governing AI systems. Interest in participatory approaches for AI has been sparked by the demonstrated success of participation methodologies in other domains of technology design, across domains such as HCI [41], education [4], international development [12], and environmental justice [45]. In addition, the potential for AI systems to inflict harm, particularly discriminatory impacts, is well-documented [10, 19, 26, 51]. There is also a growing recognition that AI systems make decisions and govern. In democratic societies, governance technologies require more democratic design and oversight processes to be acceptable and accountable [48, 57]. Hence, AI researchers are trying to identify practices and methods to ensure automated technologies serve and represent the public by expanding participation in designing, implementing, and auditing autonomous systems.

However, public participation is not a standalone solution to mitigate the harm AI systems can pose [53], and the relationship between participation and trustworthiness of AI systems is complex and unproven. Literature has highlighted 'participation washing' - a practice in which companies use participation as a performance of inclusion - illustrating the potential pitfalls and extractive practices these methods can have [8, 53]. Furthermore, a significant degree of conceptual ambiguity exists concerning what participation in AI entails and which approaches should be taken [23]. The methods, theoretical frameworks, and the stakeholders engaged matter deeply to whether participation can be a democratizing and trust-enhancing approach to AI design, or serve only as a bandaid (or worse, a fig leaf) for algorithmic harms.

This paper reviews the current state of global participatory AI projects, analyzing who is conducting PAI, the stakeholders they recruit, the theoretical lenses they use, and how they are innovating methods tied to AI development practices. We draw from this analysis to explore the relationship between participation, trust, and trustworthiness in AI systems. The paper begins by introducing trust challenges in AI, and the concept of participatory design in prior work. We then explain the methods which guided our systematic review of how global projects use participatory methods in AI development, drawing insights from 95 PAI projects. Our findings explore how projects in different parts of the world and in different economic sectors tend to choose different theoretical frameworks, ranging from techno-solutionism to decolonization. Our findings also demonstrate the relationship between methods largely drawn from previous forms of participatory design research and the *shared tasks* required for participatory AI development. Our discussion maps shared tasks developed for PAI - some tasks well-developed, others under-developed - to concepts of trust and trustworthiness, examining how sharing tasks may or may not impact trust in, and the trustworthiness of, AI systems. We close with recommendations for FAccT researchers to use the intersection of shared tasks, trust, and trustworthiness to guide and refine PAI approaches.

2 PARTICIPATION AND TRUST IN AI

PAI encompasses participation methods for designing systems, as well as for evaluation, auditing, and governance of those systems. advocates hope that participatory approaches in designing, deploying, and auditing AI technologies can advance societal-level goals such as fairness, inclusion, justice, accountability, and democratic values [23].

At the design and planning phase, participatory design (PD) is a methodology that incorporates the perspectives and experiences of technology stakeholders (including direct users, as well as sometimes indirect stakeholders impacted by a technology) as co-designers in the development of upcoming artifacts, projects, work practices, or interventions [49]. PD methodology is informed by the values of democracy and participation [56], and researchers and developers often interpret PD methods as ways to "moralize technology" [57] by ensuring that artifacts align with the values and visions of the people they will impact. Participatory Design as a methodology originated in Scandinavia in the early 1970s, primarily to democratize workplace technology [22]. Projects were often driven by trade unions or collaborations between unions and workplace researchers to empower workers as IT systems were introduced into their workplaces [56]. Early practitioners countered traditional methods that overlooked workers' tacit knowledge by using "cooperative prototyping" techniques to produce shared artifacts between designers and practitioners. This led to a "third space" for relationship building and mutual learning, fostering an inclusive design process [41].

When done well, PD can address power and knowledge imbalances in technology development by integrating diverse forms of expertise into design settings [24]. However, the degree of public participation in technology development practices can vary widely. One of the most referenced typologies is Sherry Arnstein's “Ladder of Citizen Participation” [2], which outlines degrees of participation in public planning. The ladder ranges from non-participatory methods at the bottom to 'citizen control' at the top, with various degrees of consultation, partnership, and dialogic methods in between [2, 8].

Degrees of participation have been used to map levels of participation within AI design methods such as crowdsourcing, participatory dataset documentation, creating 'red teams' for model testing, bug bounties, and public involvement in algorithmic design [23]. However, early participatory methods for AI have frequently prioritized participant quantity over the depth or length of participant involvement, leading some scholars to propose alternative approaches, such as Community Based System Dynamics (CBSD), that center marginalized and vulnerable communities' perspectives [23]. For Arnstein and researchers using the ladder, participation without the distribution of power maintains the status quo, becomes extractive, and facilitates an unjust exercise of authority [2, 8]. Bratteteig and Wagner [9] expand such critique by insisting designers must relinquish some of their power to genuinely collaborate with users as co-designers, acknowledging users' different but equally valuable expertise. Researchers like Sloane et. al [53] and Groves et al. [23] warn of participation being used to "ethically wash" an institution's practices, serving as a superficial stamp of approval. Furthermore, there are instances in which actors aim solely to improve a product [7] and include only direct users in their processes but not other impacted stakeholders, disregarding the economic, political, and social context in which technology is developed and used [53]. As a result, participation becomes exploitative and undermines co-design's potential as a truly collaborative and empowering process.

These challenges underscore the importance of a critical and nuanced understanding of participatory design to enable its practical application in AI development. However, the broadness of the term means there is no unanimous understanding of what participation in AI is, who it should serve, how it should be employed in specific contexts, or how it relates to existing design, evaluation, or governance mechanisms [8].

Resonant with, but not fully intersecting with, discussions of participatory and democratic AI have been discussions of *trustworthy* AI and *trust* in AI [58]. Trustworthy AI is a term adopted by researchers to define parameters for artificial intelligence systems that are safe, reliable, and acceptable to their users and a broader public [52]. Trustworthy AI is a design and regulation goal: building and maintaining systems that deserve the trust of the public.

Trust (a property of people) is slightly distinct from trustworthiness (a property of a system). Trust in technology is generally understood to indicate willingness to adopt and use automated systems [25], as in Lee and See's [33] early definition of trust as "the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability" (p. 54). However, measuring and fostering trust in automated systems is a complex topic, as trust is not a directly measurable feature. Instead, it is a construct made up of individual, institutional, situational and cultural factors. For example, a systematic literature review of research on trust in automation [25] grouped factors considered within trust in automation into three categories: dispositional trust, situational trust, and learned trust. Dispositional trust is reliant on individual factors such as culture, age, gender and personality, and will be variable across people. Situational trust is a category made up of factors both external (e.g. task difficulty, perceived risks, organizational setting) and internal (e.g. self-confidence, expertise, mood) but is most strongly influenced by the environment and context of an interaction. Learned trust relies on past interactions relevant to a current situation and takes into account both a user's preexisting knowledge as well as the system's performance (and how designed features of a system affect and interpret that performance). Within each of those categories, there's a reliance on competence and motivation as signals.

In this paper, we will explore how participation enables opportunities to build *trusted systems*, as defined and operationalized by Lee and See [33] and Hoff and Bashir [25]. We will also explore how participation enables opportunities to develop *trustworthy* systems by incorporating participant perspectives about what is safe, reliable, and acceptable. The next section will explain how we conducted a systematic review of global PAI projects to understand how they define and operationalize participation.

3 METHOD

This section explains our approach to the systematic review, detailing our positionality, the search methodology, the resulting dataset, and our analysis.

3.1 Identifying Participatory AI Projects

Our research is a collaboration between a Colombian graduate student conducting research in the United States and an American faculty member funded by a research institute interested in PAI. Our broad goal is to understand ways in which participation in AI development, deployment and governance does or does not support trustworthy and responsible AI. To begin work on this objective, we asked the following research questions:

RQ1. What are the theoretical frameworks being used by global PAI projects?

RQ2. What methods are being used by global PAI projects? In what ways are global PAI projects developing shared tasks between diverse stakeholders and AI experts?

RQ3. How do these theoretical frameworks, methods, and shared tasks support trust or trustworthy AI?

To identify participatory AI projects to include, we relied on Google Scholar, Google Search, and YouTube Search. Our intention to explore the PAI landscape beyond academic publications shaped our choice to utilize broad search tools, as we understand many PAI projects may not include academic publication as a goal. While Google Scholar and Google Search enabled the discovery of diverse content types such as reports, white papers, presentations, and interviews, YouTube facilitated access to academic presentations, conference talks, and project presentations spanning sectors and regions.

To ensure inclusivity, we employed keywords in both English and Spanish. This linguistic diversity aimed to identify projects not translated into English, focusing on specific communities, especially in Latin America. Search combinations like "AI+participacion+investigacion+Latam," and "AI+Inclusive+Africa+Research" or "Participatory Ai+ Inclusion+Design" enabled us to unearth numerous projects. However, for Asian and African regions, our study was constrained to English, marking a limitation of this study. We further restricted our results to publications from 2017-2024 to ensure currency. We conducted our initial searches in summer 2023, and updated them in summer 2024. We used Zotero to organize, read, and annotate pertinent content.

From our search results, we investigated all projects that included direct discussions about participatory AI in design, deployment, evaluation, governance, or research. We ascertained relevance by reading abstracts and methods sections to ensure that chosen projects engaged stakeholders beyond a core research team to develop, design, evaluate, audit, or govern an algorithmic system. We discarded projects that failed to show clear stakeholder engagement. Excluded projects included purely theoretical papers or projects that evaluated ML systems but did not engage with stakeholders beyond the research team. We included projects that explained how they had engaged with stakeholders to improve their ML/AI system, to what ends, and through which methods. Our final set comprised 95 projects.

In most cases, the projects we included self-identified using the term “participatory” within project materials. However, we did include a few projects where the methods and objectives showcased an inclusive approach with broad stakeholders, even if the authors did not explicitly use the word "participatory." Such additional cases came largely from HCI literature (where participatory methods exist as the end of a spectrum that includes co-design and user-centered design), such as developing AI technology for the visually impaired, conflict resolution, or education enhancement for children.

The aim of our review was not to create an exhaustive dataset of global participatory AI projects. Comprehensiveness is a longer-term goal that will be pursued in future work through additional methods such as crowdsourcing and citation chaining. Instead, we focused on creating a diverse initial database of exemplar projects: projects that added new geographic scope, theoretical lenses, forms of participation, or methods to the database. We stopped searching when we no longer found projects that added one of these new features. Our final dataset demonstrates the diversity of how different AI projects across the globe engage with users, other actors, stakeholders, or citizens through diverse participatory methods.

Our next step was to use the database to answer RQs 1 and 2, investigating the theoretical frameworks and methods currently used by represented projects. We read or watched each source document and manually categorized them according to a codebook we designed to help us understand how and why PAI was being used, as explained in Table 1. At this point, we moved out of Zotero and created an AirTable database, which facilitated the categorization and analysis of projects by categories of interest (Appendix 1). We developed categories to match our project's interest in geography, sector, theoretical and methodological approaches, and definitions of participation. Geography categorization included regions (North America, Latin America, Europe, Asia, Oceania, and Africa) and countries. Sectors categorized project organizers by domain (for example, NGO, Academic Partnership, Public Institutions, Government, Private Companies, Multi-Sector Coalition, and Multi-Actor Collaboration).

Theoretical lenses comprised the assumptions, motivations, and ideas employed by researchers or actors, which guided approaches and informed methods [21]. We identified those lenses by analyzing introductions, justification sections, and positionality statements when available. Some projects clearly stated their lens; in other cases, we inferred lenses from project descriptions. We identified 17 different types of theoretical lenses in our dataset [see appendix 2 for the full list]. Some of the lenses we identified include civic engagement (in which stakeholders are encouraged to engage in policymaking [27]); participatory feminism (a critical approach that explores gender-based systemic biases in tech [16, 20]); data justice (a framework that examines data issues in the context of existing power dynamics, ideology, and social practices [55]); techno-solutionism (approach that prioritizes solving problems through technological innovation [39]); and decolonization (an approach that centered non-Western standpoints and power [38]).

Finally, we described the approaches used in the project to engage participants as methods. We found methods by reading project descriptions and method sections, where available. A complete list is included in Appendix 2, but methods included alliance creation, focus groups, role-play simulation, observations, ethnography, cooperative inquiry, and red teaming. For projects with more detailed methods descriptions (57 of the 95 projects), we further mapped the methods to the development lifecycle for both machine learning classifiers and generative AI. This mapping helped us identify the relationship between existing methods in the field and the shared tasks necessary to support AI development.

Several projects in the database included more than one method, theoretical lens, or actor type. In order to highlight such diversity and overlapping, our coding allowed for more than one code in each category, and our analysis took overlapping codes into account. Organizing projects by multiple categories helped us understand what lenses and methods are well represented in PAI research, as well as which are not. We discuss these findings in the next section.

4 FINDINGS

This section describes the theoretical lenses and engagement methods discovered in our dataset. We also present trends by economic sector, discussing characteristics of participatory AI led by government organizations, NGOs, private companies, commercial AI labs, and academic research institutions. Finally, we describe how current participatory methods map to AI development tasks.

Our final dataset consisted of 95 projects that showcased participation methods for creating, designing, deploying, evaluating, and governing machine learning and artificial intelligence systems. The final dataset exhibits a notable skew towards the Global North, with North American and European projects comprising 52% of the sample. In contrast, Asia (8), Africa (7), Latin America (6), and Oceania (4) collectively contribute fewer projects, illustrating a well-established challenge of identifying and highlighting Global South initiatives within the predominantly Global North-dominated AI landscape. Our use of only English and Spanish search terms may have contributed to this bias, presenting a limitation while offering an opportunity for future research. Additionally, 14 projects are developed by partners in more than one region, indicating collaboration among actors across diverse geographical areas.

Distinct regional and thematic divisions are evident, particularly in the theoretical lenses adopted, as shown in Figure 1. As discussed in Section 3, a theoretical framework comprises concepts, definitions, and existing theories, connecting the study to broader knowledge. It shows the assumptions, motivations, and ideas employed by project organizers, which guide approaches and inform methods [21].

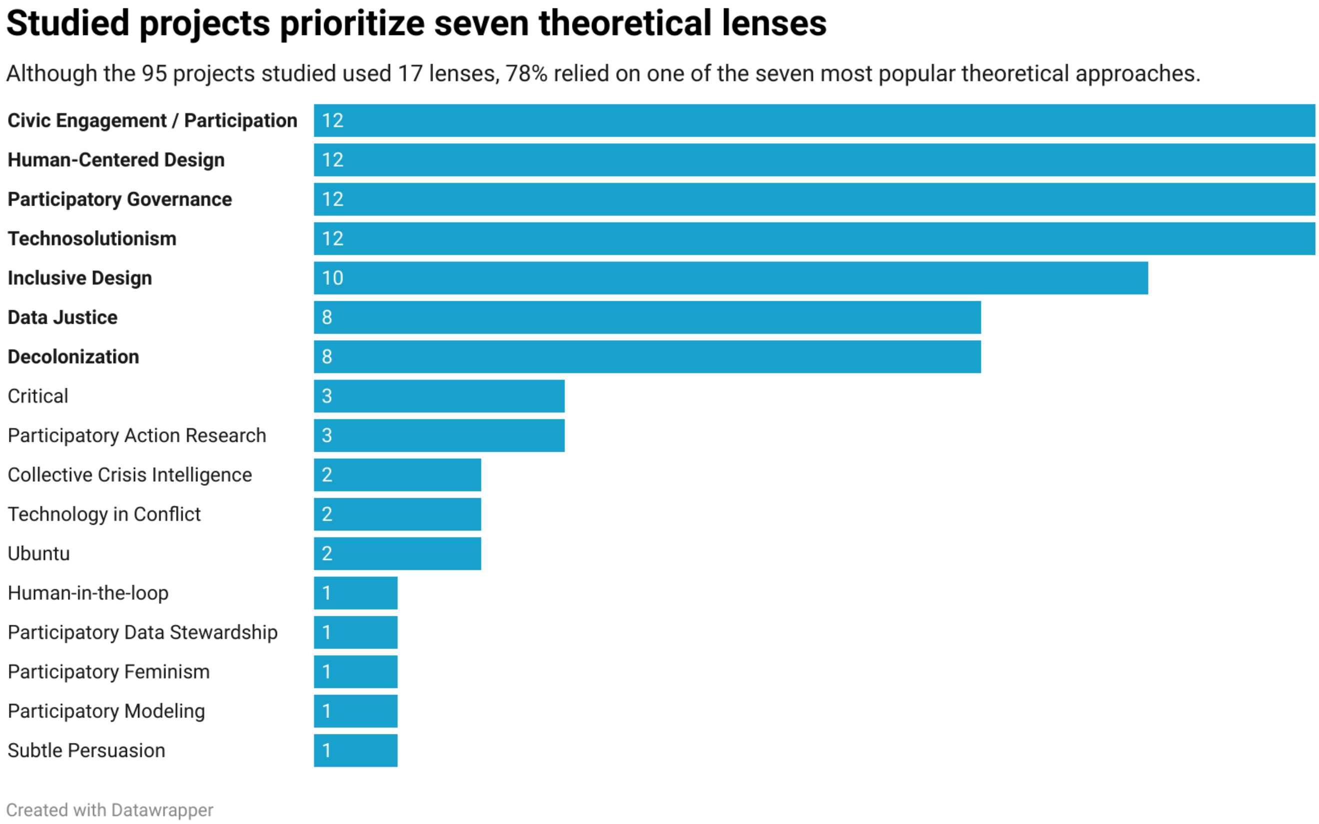

Figure 1: Bar chart showcasing the different lenses identified in the database and the number of projects per lens

Participatory governance, civic engagement, and human-centered design were the most common theoretical framework across the dataset, encompassing 38% of the projects studied. This underscores international governments' current emphasis on participation within approaches to regulating AI, involving the design of national AI strategies, frameworks, and sandboxes. This trend transcends regional boundaries, indicating a global convergence toward addressing AI's ethical, political, and societal implications through regulation.

Outside of the dominance of participatory governance, we found regional differences in theoretical approaches to participatory AI. For instance, participatory feminism, a critical approach focusing on addressing gender-based systemic biases, was identified solely in projects based in Latin America. Projects centered in Africa tended to emphasize decolonization and data justice lenses, highlighting a commitment to rebalancing power dynamics and knowledge structures within AI. In contrast, projects from the Global North favor frameworks such as human-centered design, inclusive design, or techno-solutionism, prioritizing problem solving through technological efficiency and a market-centric approach to social change.

Despite the diversity in theoretical frameworks, project methods in the dataset overall demonstrate adoption and adaption of qualitative approaches frequently used in other forms of community-based research, as shown in Figure 2. A plurality of projects utilize working groups or workshops (14%) and interviews (12%). Some projects also added methods to gather data from stakeholders at larger scales, such as crowdsourcing (7%) and surveys (4%). The dominance of qualitative methods persists irrespective of the theoretical framework. For instance, among the 19 projects guided by an inclusive design theoretical framework, 68% incorporate at least one of the qualitative approaches mentioned. Of the 22 projects with a human-centered design framework, 60% use at least one of these methods, as do 47% of the ones adopting a techno-solutionism framework, and 45% of the projects using decolonization frameworks.

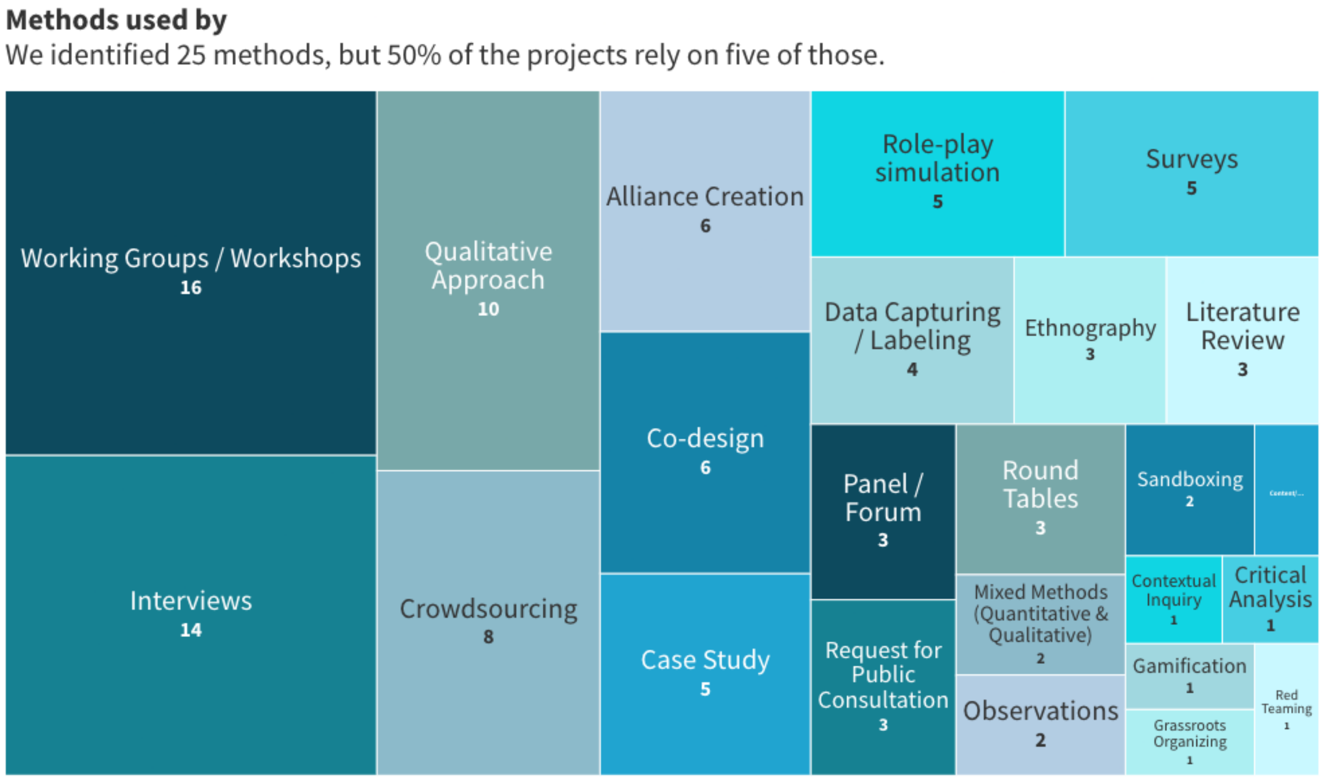

Figure 2: Treemap showcasing the identified methods and the number of projects per method.

4.1 Sectoral differences in approaches

In this section, we discuss patterns in how particular sectors engaged PAI lenses and methods.

Government Agencies: Published research, white papers, and official documents found in our search show some national governments worldwide are turning to participatory approaches in AI development and regulation. Government agencies employ both participatory design and participatory governance methods. For example, projects in our results from Canada, Colombia, Mexico, and Spain show the diversity of participation methods and lenses governments are applying. The Canadian government deployed five strategies to incorporate public participation in the design AI policy: working groups (groups of diverse stakeholders that get together over time to discuss and solve specific issues) focused on public policy and AI justice; entrepreneur meetings; workshops with companies; public comment periods; and seminars with activists, NGOs, and citizens. In contrast, the Colombian strategy focused solely on public comment periods and small working groups with national and international experts. Mexico conducted a national survey that informed working groups with local experts and critical institutions. Spain was the first country in the European Union to test the EU approach to AI regulation through a sandbox, a temporary testing space to experiment, enabling trial-and-error processes to assess the adequacy of existing legal frameworks. We also noted that, between our initial searches in 2023 and follow-up searches in 2024, information about projects in several countries disappeared following changes in government leadership. For instance, Colombia's AI policy website was deactivated when a new administration came to power. The absence of previously participatory AI strategies highlights a challenge for participatory governance efforts as political changes shift government approaches.

NGOs and NPOS: Our review illustrates that the vast majority of grassroots organizations, NPOs, and NGOs actively champion theoretical lenses like data justice, decolonization, and civic engagement within participatory AI projects. These lenses prioritize dismantling power imbalances, ensuring equitable access to AI benefits, and promoting meaningful public engagement in AI development and governance, especially in challenging circumstances due to poverty or conflict. Their methods echo these priorities, heavily relying on collaborative techniques like interviews, working groups, workshops, alliance creation, and crowdsourcing, often in conjunction with co-design approaches. The Masakhane MT project exemplifies this participatory spirit, where community members drive machine translation development for less-resourced languages through crowdsourcing and co-design [36]. Similarly, EmpatIA, a Latin American project led by the civil society organization ILDA, utilizes mixed methods to engage citizens in auditing public and private AI adoption, ensuring the ethical, political, social, and economic considerations of AI development are embedded in public policy [18]. Another interesting example is the project WeBuildAI, an alliance between 412 Food Rescue and the University of Texas [34]. By using interviews, working groups, and pairwise ranking, researchers and volunteers created a matching algorithm that operates an on-demand food donation transportation service to adjudicate equity and efficiency trade-offs. These grassroots efforts showcase civil society's critical role in shaping participatory methods for AI.

Academia: Our analysis found that universities and research institutions engaged in PAI are most likely to use approaches developed in human-computer interaction (human-centered design and inclusive design) as well as broader software development practices such as agile design. Theoretical lenses and methodological choices common in academic projects reflect roots in HCI and software research. Projects like the AuXie Blind-Accessible Virtual Museum Tour [17] and the MYCam teachable object recognizer app for blind users [28] center end users in design, testing, seeking to ensure the resulting systems are functional but also reliable and trustworthy for users. Alongside NGOs, universities and research institutions play a crucial role in pushing the boundaries of AI development and participation, especially as researchers forge alliances with actors who may lack the resources, funding, and expertise to develop AI projects independently.

Private companies: Some private companies and commercial AI labs are also exploring participatory AI. Though many corporate projects have a distinct emphasis on techno-solutionism, which aligns with traditional entrepreneurial views of technology [42], a few of the most critical projects have also been conducted by teams within corporations [23]. Corporate techno-solutionist projects can also develop innovative participation methodologies, such as the ones used in SynthBio and The Prompt Artist. In Google's SynthBio project, AI supports a novel collaboration for efficient dataset curation. Researchers use a large language model to provide seed generations for humans performing crowdsourced data labeling, changing crowdsourced data labeling into an editing task [59]. Meanwhile, Google researchers working on the Prompt Artist project created a dialogue between researchers and AI-created artists to better understand the polemic practice of AI art creation [11]. Both projects are examples of the techno-solutionist spirit, where active participation accelerates creativity in technical innovation.

4.2 From participatory methods to shared tasks

Comparing the methods employed in exemplar projects to the AI development pipeline illustrated the considerable progress the field of PAI has made in adapting participatory methods from other fields for AI development. To map the AI pipeline, we turned to Daume's "A Course in Machine Learning" [13], expanding what he describes as a "typical design process for a machine learning application" to also include newer processes for considering, adopting, and adapting large language and large image models. This expansion led to using the following ML pipelines to describe tasks required in AI development:

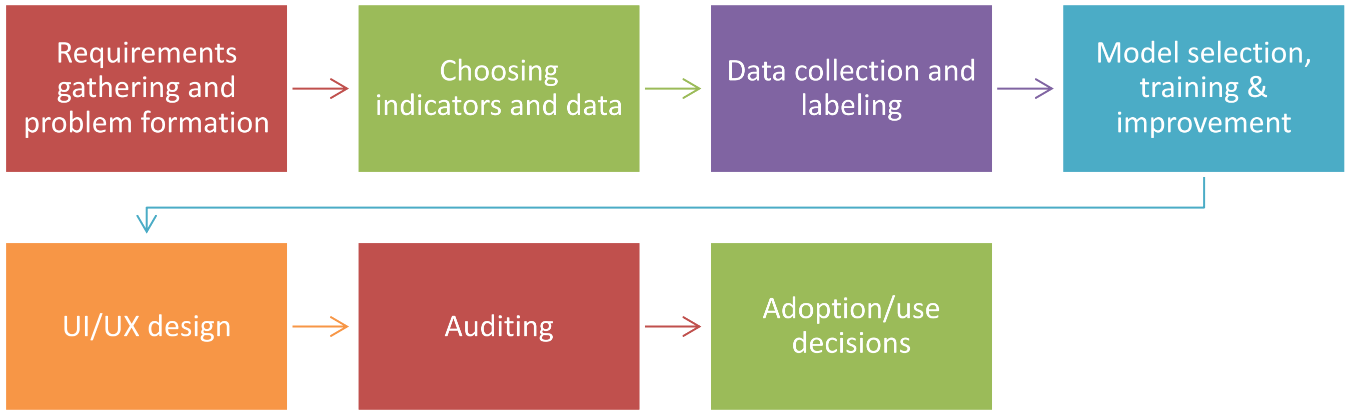

Figure 3: Process representation of the traditional ML lifecycle

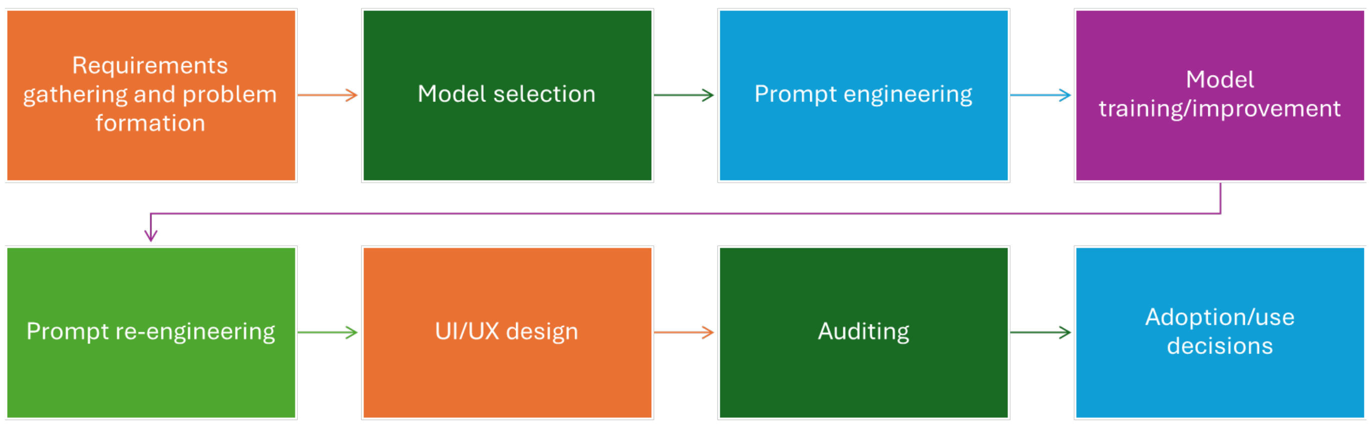

Figure 4: Process representation of an LLM lifecycle

Several projects in our database developed participatory methods for the first step in the machine learning lifecycle: ***requirements gathering and problem formulation***, or deciding why, exactly, an AI technology is needed and what challenges it should address. Examples of participatory requirements gathering included gathering vignettes and counternarratives from indigenous stakeholders in a project focused on Indigenous AI [35]; creating community values frameworks to guide design, such as the Lakota Good Way framework [35]; and conducting role-playing and simulation exercises to define problems that could be addressed through automation [14]. The WeBuildAI project, mentioned above, partnered with a volunteer-based nonprofit to determine needs, goals, and ways of specifying those goals [34]. A research team also partnered with civil rights nonprofits used contextual inquiry, an HCI technique, to do requirements gathering to guide creation of the Algorithmic Equity Toolkit [32].

Examples of participatory methods for the next step in the AI lifecycle, ***choosing both indicators and appropriate data to train classifiers****,* include a co-design process conducted through role playing and simulation, which led to cross-disciplinary expertise sharing in the TREC legal discovery project [14]. Communities also collaborated to crowdsource appropriate data to train classifiers for machine translation of African languages [8]. Participatory methods for conducting data collection and labeling for classifiers include participatory data stewardship techniques documented by the Ada Lovelace Institute [60]. The Open for Good project encourages data donation as a means to ensure participation in data sourcing [29].

Examples of participatory methods for ***model selection, training and improvement*** include several techniques developed by the project WeBuildAI [34]. These include having stakeholders assign scores to tell the algorithm how to weigh particular features, and later visualizing models and allowing participants to select the one that best expresses their beliefs and preferences. The Nesta group [5] provides multiple case studies in participation to improve model performance, such as Human-in-the-Loop Project Dorian, which involved first responders labeling data from social media about Hurricane Dorian, and then assessing the accuracy of a learning classifier in real time. The Nesta group also describes Project Sepsis, which (among a variety of participatory processes) had clinicians evaluate the accuracy of a deep learning model in identifying known sepsis cases, and then tuned the model based on clinician feedback [5].

Searching for participatory ***interaction design for LLMs*** revealed a gap in the current literature. Prompt engineering currently dominates interaction design for LLMs (although that may not always be the case [40]). Although on some level, prompt engineering could be defined as already participatory (as it is conducted by stakeholders using, not building, AI systems), organized efforts to study community sense-making involved in prompt selection or other interaction designs may be needed to understand ways in which prompts shape results from, and interactions with, large language models.

Examples of participatory ***interface design*** for AI system interfaces include work in Project Sepsis [5], in which nurses engaged in participatory design of the interface. The area of explainable AI has also produced work on participatory design of interfaces, such as interviews and workshops conducted to support the design of AI for kidney transplantation [54]. There are also examples of participatory AI interface design work in the cultural heritage space, in which interfaces have been codesigned with, for example, museum visitors [46].

Examples of participatory ***auditing*** for AI systems include the Algorithmic Equity Toolkit [32], developed as part of a larger participatory design project which also included needs elicitation and prototype evaluation. The project accomplished participatory auditing by co-designing a reflexive tool meant for advocates to use to interrogate possible social and technical problems in algorithmic systems, and to create policy positions around those systems. And the US government hosted a large experiment in participatory auditing at the Defcon security conference in 2023, when it invited security experts in attendance as well as diverse invited participants to try to purposefully produce errors and better understand the risks of large language models [37].

There is also work in participatory AI on the ways that stakeholder values impact ***adoption and adaptation of AI systems***. For example, the STELA process [6] works with stakeholder communities to elicit community values through deliberative discussion, with the goal of aiding adoption systems by aligning LLM systems with those values. The nonprofit NESTA [5] highlights several case studies of projects which take user values into account to adopt and adapt AI technologies. One example is Wikipedia's Objective Revision Evaluation Service (ORES) project, in which subcommunities on Wikipedia work with machine learning engineers to tailor automated content moderation algorithms to their community's needs and specifications. Community audits then make sure that deployed classifiers align with community values.

Finally, numerous projects documented the various and diverse forms of non-technical work needed for successful participatory AI - what we might call ***participatory culture building***. Non-technical work important to participatory AI included steps like documenting labor [31], welcoming and supporting marginalized participants [47], critiquing exploitative participation practices [31, 47, 53], and providing guidance and making non-technical adjustments to situations to ensure equity and accountability [32].

5 DISCUSSION: LINKING PARTICIPATION AND TRUST

Very little literature has examined the practical relationship between participatory AI methods, as explored in depth in our findings, and components of trust and trustworthiness. Returning to the many factors which impact trust in automated systems [25], we explore here the ways that the shared tasks developed in our findings could support and scaffold components of trust and trustworthiness. We find that participatory processes can help support components of both ***situational and learned trust*** – the external factors that influence trust in automated systems – by helping participants better understand risks of AI systems, organizational norms impacted by automation, and details of system performance. We also find that participatory process might support internal factors that contribute to situational trust by building participant self-confidence, expertise, and preexisting knowledge. However, relying on participation alone to build trust in systems is an expensive proposition: though direct participants may end up with increased learned and situational trust in systems, such granular trust-building may only indirectly transfer to broader populations. More important to participatory AI is that participatory processes support *trustworthy* AI by informing the system components and functions defined as safe, functional and acceptable.

Requirements gathering and problem formation	Trust elements: Knowledge of impact - Participants understand why a problem was selected for automation.
	Trustworthy elements: Values - Developers choose a problem well-suited to the context.
Choosing data or indicators	Trust elements: Knowledge of performance - Participants understand why training data or indicators were selected.
	Trustworthy elements: Developers choose data and indicators better informed by the context.
Data collection and labeling	Trust elements: Knowledge of risks – Participants understand what data is included and excluded (or unavailable).
	Trustworthy elements: Functionality, Acceptability – Developers able to spot bias and unacceptable data uses.
Model selection, training and improvement:	Trust elements: Knowledge of risks and performance – Participants understand what a model optimizes.
	Trustworthy elements: Functionality, Acceptability – Developers choose and tune appropriate models for the context.
Prompt engineering	Trust elements: Knowledge of performance – Participants understand what impacts prompts have on outputs.
	Trustworthy elements: Functionality, Safety – Developers can standardize interaction techniques.
UI/UX design	Trust elements: Knowledge of performance – Participants understand what system outputs do and do not represent.
	Trustworthy elements: Acceptability – Developers include representations that stakeholders need to interpret outputs.
Auditing	Trust elements: Knowledge of risks and performance – Participants understand frequency/nature of problematic outputs.
	Trustworthy elements: Safety – Developers gain feedback on unacceptable outputs.
Adoption decisions	Trust elements: Knowledge of risks, performance, and impacted organizational norms.

Figure 5: Shared tasks, trust, and trustworthiness

Figure 5 outlines granular ways that shared tasks described in global participatory AI projects contribute to elements of trust and trustworthiness. Participatory techniques for AI requirements gathering, problem formulation, and interface design are fairly well-developed in the literature we have reviewed, and in many ways, echo forms of trust and trustworthiness that have long been attributed to participatory design. But for shared tasks that are specific to AI development, such as choosing indicators and appropriate training data; model selection, training and improvement; and interaction engineering, our review of the literature demonstrates new ways in which participatory processes can contribute to both trust and trustworthiness. For example, work on participatory ways to choose indicators and appropriate training data, such as that by WeBuildAI [34] or the M¯aori Data Sovereignty Protocols [35], bolsters trust by providing participants with deep situational and learned knowledge about the fit of indicators and data on which outputs will rely. And participatory indicator and training data selection simultaneously bolsters trustworthiness by guiding developers on the appropriateness of particular datasets for system design. Similarly, participation in the auditing and evaluation of AI system outcomes, such as that explored by the AI Red Teaming Challenge (cite) improves participants' situational and learned knowledge of AI outputs, errors, and bias. It can also improve the trustworthiness of systems if developers can address and mitigate the errors and biases red teaming participants find. Participation in shared tasks like choosing training data, model selection, and interaction engineering can help avoid "abstraction traps" as identified by Selbst et al [50] by incorporating explicitly social knowledge from participants into development processes.

However, the dominance of technological-solutions lenses in the Global North found in our analysis raises concerns about the neglect of power dynamics and social justice considerations explored in other regions through critical lenses, such as data justice, decolonization, or participatory feminism. Our review of theoretical frameworks across economic sectors also points to the importance of participatory AI work happening not only in industry, academia, and government, but also by nonprofits and NGOs. In many cases, nonprofits and NGOs have not only deep connections with stakeholder communities, but also the incentives to frame work with truly emancipatory participatory lenses. We found that some of the most innovative participatory AI projects, with the most concrete interventions into shared AI tasks, came from NGOs.

If trustworthiness (and indirectly, trust) is to be supported through participatory methods in AI, our analysis underscores the necessity of project reporting, whether YouTube videos, white papers, or academic publications, providing comprehensive descriptions of participants' involvement, power dynamics, and design decisions incorporating shared tasks. To unlock the full potential of participation in advancing AI, there must be a concerted effort to establish improved protocols for reporting on the participation process and its relationship to trustworthy AI. Only through these measures can PAI genuinely realize its potential for driving meaningful impact in artificial intelligence.

6 CONCLUSION

Participatory AI (PAI) holds immense promise for fostering trustworthy development and ensuring societal benefit from artificial intelligence's increasingly pervasive influence. However, our analysis of current projects across diverse actors unveils a nascent a field with complexities and contradictions. The promise of PAI to foster trust in automated systems is still in development, allowing ample space to innovate and improve. Our analysis reveals a field fragmented by diverse and ample definitions of participation, with varying degrees of interaction and co-creation and diverse levels of inclusion of actors, citizens, and stakeholders. Most projects' methodologies borrow from established participatory design practices, without significant innovation tailored to AI's unique challenges. Future PAI efforts should encompass the design, auditing, implementation, and governance of algorithmic systems. Such innovation is necessary for PAI to live up to its promise of equitable and trustworthy AI development.

7 REFERENCES

[1] Afroogh, S., Akbari, A., Malone, E., Kargar, M. and Alambeigi, H. 2024. Trust in AI: progress, challenges, and future directions. *Humanities and Social Sciences Communications*. 11, 1 (Nov. 2024), 1–30. DOI:https://doi.org/10.1057/s41599-024-04044-8.

[2] Arnstein, S.R. 1969. A ladder of citizen participation. *JAIP*. 35, 4 (Jul. 1969), 216–224.

[3] Bender, E.M. and Friedman, B. 2018. Data Statements for Natural Language Processing: Toward Mitigating System Bias and Enabling Better Science. *Transactions of the Association for Computational Linguistics*. 6, (Dec. 2018), 587–604. DOI:https://doi.org/10.1162/tacl_a_00041.

[4] Benton, L. and Johnson, H. 2015. Widening participation in technology design: A review of the involvement of children with special educational needs and disabilities. *International Journal of Child-Computer Interaction*. 3–4, (Jan. 2015), 23–40. DOI:https://doi.org/10.1016/j.ijcci.2015.07.001.

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

[6] Bergman, S., Marchal, N., Mellor, J., Mohamed, S., Gabriel, I. and Isaac, W. 2024. STELA: a community-centred approach to norm elicitation for AI alignment. *Scientific Reports*. 14, 1 (Mar. 2024), 6616. DOI:https://doi.org/10.1038/s41598-024-56648-4.

[7] Bergvall-Kåreborn, B. and Ståhlbrost, A. 2008. Participatory design: one step back or two steps forward? *Proceedings of the Tenth Anniversary Conference on Participatory Design 2008* (USA, Oct. 2008), 102–111.

[8] Birhane, A., Isaac, W., Prabhakaran, V., Díaz, M., Elish, M.C., Gabriel, I. and Mohamed, S. 2022. Power to the People? Opportunities and Challenges for Participatory AI. *Equity and Access in Algorithms, Mechanisms, and Optimization* (Washington D.C., Oct. 2022), 1–8.

[9] Bratteteig, T. and Wagner, I. 2016. Unpacking the Notion of Participation in Participatory Design. *Computer Supported Cooperative Work (CSCW)*. 25, 6 (Dec. 2016), 425–475. DOI:https://doi.org/10.1007/s10606-016-9259-4.

[10] Buolamwini, J. and Gebru, T. 2018. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Conference on Fairness, Accountability and Transparency* (Jan. 2018), 77–91.

[11] Chang, M., Druga, S., Fiannaca, A.J., Vergani, P., Kulkarni, C., Cai, C.J. and Terry, M. 2023. The Prompt Artists. *Proceedings of the 15th Conference on Creativity and Cognition* (New York, NY, USA, Jun. 2023), 75–87.

[12] Cleaver, F. 1999. Paradoxes of participation: questioning participatory approaches to development. *Journal of International Development*. 11, 4 (1999), 597–612. DOI:https://doi.org/10.1002/(SICI)1099-1328(199906)11:4<597::AID-JID610>3.0.CO;2-Q.

[13] Daumé III, H. 2017. *A course in machine learning*.

[14] Delgado, F., Barocas, S. and Levy, K. 2022. An Uncommon Task: Participatory Design in Legal AI. *Proceedings of the ACM on Human-Computer Interaction*. 6, CSCW1 (Mar. 2022), 1–23. DOI:https://doi.org/10.1145/3512898.

[15] Demystifying the Black Box That Is AI: *https://www.scientificamerican.com/article/demystifying-the-black-box-that-is-ai/*. Accessed: 2020-06-29.

[16] D'Ignazio, C. and Klein, L.F. 2020. *Data Feminism*. The MIT Press.

[17] Dulyan, A. and Edmonds, E. 2010. AUXie: initial evaluation of a blind-accessible virtual museum tour. *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction* (New York, NY, USA, Nov. 2010), 272–275.

[18] Empatía | Inteligencia Artificial para el desarrollo en América Latina y el Caribe: *https://www.empatia.la/en/home/*. Accessed: 2025-01-17.

[19] Eubanks, V. 2017. *Automating inequality: how high-tech tools profile, police, and punish the poor*. St. Martin's Press.

[20] Gervais, M., Weber, S. and Caron, C. 2018. *Guide to Participatory Feminist Research*. Institute for Gender, Sexuality, and Feminist Studies.

[21] Grant, C. and Osanloo, A. 2014. Understanding, Selecting, and Integrating a Theoretical Framework in Dissertation Research: Creating the Blueprint for Your "House." *Administrative Issues Journal: Connecting Education, Practice, and Research*. 4, 2 (2014), 12–26.

[22] Gregory, J. 2003. Scandinavian Approaches to Participatory Design. *International Journal of Engineering Education*. 19, 1 (2003), 62–74.

[23] Groves, L., Peppin, A., Strait, A. and Brennan, J. 2023. Going public: the role of public participation approaches in commercial AI labs. arXiv.
[24] Harrington, C., Erete, S. and Piper, A.M. 2019. Deconstructing Community-Based Collaborative Design: Towards More Equitable Participatory Design Engagements. *Proceedings of the ACM on Human-Computer Interaction*. 3, CSCW (Nov. 2019), 216:1-216:25. DOI:https://doi.org/10.1145/3359318.
[25] Hoff, K.A. and Bashir, M. 2015. Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust. *Human Factors*. 57, 3 (May 2015), 407–434. DOI:https://doi.org/10.1177/0018720814547570.
[26] Hoffmann, A.L. 2019. Where fairness fails: data, algorithms, and the limits of antidiscrimination discourse. *Information, Communication & Society*. 22, 7 (Jun. 2019), 900–915. DOI:https://doi.org/10.1080/1369118X.2019.1573912.
[27] Holley, K. 2016. *Principles for Equitable and Inclusive Civic Engagement: A Transformative Guide*. Kirwan Institute.
[28] Hong, J., Gandhi, J., Mensah, E.E., Zeraati, F.Z., Jarjue, E., Lee, K. and Kacorri, H. 2022. Blind Users Accessing Their Training Images in Teachable Object Recognizers. *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility* (New York, NY, USA, Oct. 2022), 1–18.
[29] Inclusive AI Alliance: *https://openforgood.info/*. Accessed: 2025-01-17.
[30] Jacobs, A.Z. and Wallach, H. 2021. Measurement and Fairness. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (New York, NY, USA, Mar. 2021), 375–385.
[31] Jing, F.S., Berger, S.E. and Becerra Sandoval, J.C. 2023. Towards Labor Transparency in Situated Computational Systems Impact Research. *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (New York, NY, USA, Jun. 2023), 1026–1037.
[32] Katell, M., Young, M., Dailey, D., Herman, B., Guetler, V., Tam, A., Bintz, C., Raz, D. and Krafft, P.M. 2020. Toward situated interventions for algorithmic equity: lessons from the field. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (New York, NY, USA, Jan. 2020), 45–55.
[33] Lee, J.D. and See, K.A. 2004. Trust in Automation: Designing for Appropriate Reliance. *Human Factors*. 46, 1 (Mar. 2004), 50–80. DOI:https://doi.org/10.1518/hfes.46.1.50_30392.
[34] Lee, M.K., Kusbit, D., Kahng, A., Kim, J.T., Yuan, X., Chan, A., See, D., Noothigattu, R., Lee, S., Psomas, A. and Procaccia, A.D. 2019. WeBuildAI: Participatory Framework for Algorithmic Governance. *Proceedings of the ACM on Human-Computer Interaction*. 3, CSCW (Nov. 2019), 181:1-181:35. DOI:https://doi.org/10.1145/3359283.
[35] Lewis, J.E. ed. 2020. *Indigenous Protocol and Artificial Intelligence Position Paper*. The Initiative for Indigenous Futures and the Canadian Institute for Advanced Research (CIFAR).
[36] Masakhane: *https://www.masakhane.io/*. Accessed: 2025-01-17.
[37] Mislove, A. 2023. *Red-Teaming Large Language Models to Identify Novel AI Risks*. Office of Science and Technology Policy.
[38] Mohamed, S., Png, M.-T. and Isaac, W. 2020. Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence. *Philosophy & Technology*. 33, 4 (Dec. 2020), 659–684. DOI:https://doi.org/10.1007/s13347-020-00405-8.
[39] Morozov, E. 2013. *To Save Everything, Click Here: The Folly of Technological Solutionism*. PublicAffairs.
[40] Morris, M.R. 2024. Prompting Considered Harmful – Communications of the ACM. *Communications of the ACM*. (Oct. 2024).
[41] Muller, M. 2003. Participatory design: the third space in HCI. *Handbook of HCI*. Erlbaum.

[42] Nachtwey, O. and Seidl, T. 2024. The Solutionist Ethic and the Spirit of Digital Capitalism. *Theory, Culture & Society*. 41, 2 (Mar. 2024), 91–112. DOI:https://doi.org/10.1177/02632764231196829.
[43] National Institute of Standards and Technology, I. 2016. *NIST Risk Management Framework*.
[44] Organisation for Economic Co-operation and Development *Hiroshima Process International Code of Conduct for Organisations Developing Advanced AI Systems,*.
[45] Ottinger, G. 2013. Changing Knowledge, Local Knowledge, and Knowledge Gaps: STS Insights into Procedural Justice. *Science, Technology, & Human Values*. 38, 2 (Mar. 2013), 250–270. DOI:https://doi.org/10/f5p7z2.
[46] Pisoni, G., Díaz-Rodríguez, N., Gijlers, H. and Tonolli, L. 2021. Human-Centered Artificial Intelligence for Designing Accessible Cultural Heritage. *Applied Sciences*. 11, 2 (Jan. 2021), 870. DOI:https://doi.org/10.3390/app11020870.
[47] Queerinai, O.O. et al. 2023. Queer In AI: A Case Study in Community-Led Participatory AI. *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (New York, NY, USA, Jun. 2023), 1882–1895.
[48] Risse, M. 2021. *Artificial Intelligence and the Past, Present, and Future of Democracy*. Carr Center for Human Rights Policy.
[49] Schuler, D. and Namioka, A. 1993. *Participatory Design: Principles and Practices*. Lawrence Erlbaum Associates.
[50] Selbst, A.D., Boyd, D., Friedler, S.A., Venkatasubramanian, S. and Vertesi, J. 2019. Fairness and Abstraction in Sociotechnical Systems. *Proceedings of the Conference on Fairness, Accountability, and Transparency* (New York, NY, USA, Jan. 2019), 59–68.
[51] Shelby, R., Rismani, S., Henne, K., Moon, Aj., Rostamzadeh, N., Nicholas, P., Yilla-Akbari, N., Gallegos, J., Smart, A., Garcia, E. and Virk, G. 2023. Sociotechnical Harms of Algorithmic Systems: Scoping a Taxonomy for Harm Reduction. *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society* (New York, NY, USA, Aug. 2023), 723–741.
[52] Shneiderman, B. 2020. Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy. *International Journal of Human–Computer Interaction*. 36, 6 (Apr. 2020), 495–504. DOI:https://doi.org/10.1080/10447318.2020.1741118.
[53] Sloane, M., Moss, E., Awomolo, O. and Forlano, L. 2020. Participation is not a Design Fix for Machine Learning. *Proceedings of the 37th International Conference on Machine Learning* (Vienna, Austria, 2020).
[54] Subramanian, H.V., Canfield, C. and Shank, D.B. 2024. Designing explainable AI to improve human-AI team performance: A medical stakeholder-driven scoping review. *Artificial Intelligence in Medicine*. 149, (Mar. 2024), 102780. DOI:https://doi.org/10.1016/j.artmed.2024.102780.
[55] Taylor, L. 2017. What is data justice? The case for connecting digital rights and freedoms globally. *Big Data & Society*. 4, 2 (Dec. 2017), 2053951717736335. DOI:https://doi.org/10.1177/2053951717736335.
[56] van der Velden, M. and Mörtberg, C. 2021. Participatory Design and Design for Values. *Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains*. J. van den Hoven, P.E. Vermaas, and I. van de Poel, eds. Springer Netherlands. 1–22.
[57] Verbeek, P.-P. 2006. Materializing morality. *Science, Technology & Human Values*. 31, 3 (May 2006), 361–380.
[58] Vereschak, O., Bailly, G. and Caramiaux, B. 2021. How to Evaluate Trust in AI-Assisted Decision Making? A Survey of Empirical Methodologies. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2 (Oct. 2021), 327:1-327:39. DOI:https://doi.org/10.1145/3476068.
[59] Yuan, A., Ippolito, D., Nikolaev, V., Callison-Burch, C., Coenen, A. and Gehrmann, S. 2022. SynthBio: A Case Study in Human-AI Collaborative Curation of Text Datasets. arXiv.

[60] 2021. *Participatory data stewardship*. Ada Lovelace Institute.

8 APPENDICES

1.1 APPENDIX 1

Access link to Air Table: [blinded for review]

1.2 APPENDIX 2

Table 1. Codebook for the Categorization of Projects

Category	Definition
Region	The database encompasses a global perspective, categorizing projects based on geographical region. This allowed us to see if there were any regional trends in approaches. If the project is designed or directed by several regions, the category will state "Multi." If not, it was classified as follows: North America (Canada, United States) Latin America (Mexico, Central America and the Caribbean, and South America) Europe Asia Oceania Africa
Country	If the project is designed, implemented, or directed by several countries, the category will state "Multi." If not, it will state the country from which the project is originated.
Actor Type	Describes the type of actor that develops the projects. We found it by looking into the author's affiliation or the organization leading the project. We chose only one of the following options to avoid redundancy and better observe trends and relationships: Civil Organization / Grass Roots Community Research Network/ Partnership (Alliance between universities and research centers) Government Multi-Sector Coalition Private Companies Public Institutions

	Multi-Actor Collaboration (used for projects led by multiple actors with diverse backgrounds. i.e. governments, industry, and academia) Multilateral Collaboration (IMF, World Bank, UN, etc.) University (if project is only led by one university) NGO Academic Partnership (if project is led by diverse universities) International Organization (EU, OAS) NPO Research Center
Actor Name	Includes the actor’s name. If more than one actor, include all of them.
Actor Description	Briefly explain what the actor(s)/ institution(s) does if needed.
Project Name	Include the name of the project or research paper.
Project Description	Briefly explain what the project does and why it is PAI.
Link	Paste the link to the research project, paper, or project site.
Lenses	A lens or theoretical framework comprises concepts, definitions, and existing theories, connecting the study to broader knowledge. It shows the assumptions, motivations, and ideas employed by researchers or actors that guide approaches and inform methods (Grant & Osanloo, 2014). We identified those lenses by analyzing introductions justification sections and positionality statements when available. Some projects clearly stated their lens, while others did not. Lenses found include: ● Civic Engagement / Participation: The involvement of individual constituents or communities in local, state, and national government ● Participatory Feminism: A critical approach to studying the adoption and participation of AI by exploring gender-based systemic biases and involving participants as co-creators. ● Participatory Governance: Lens through which decision-makers involve constituents in policy-making decisions ● Data Justice: Framework that examines data issues in the context of existing power dynamics, ideology, and social practices ● Ubuntu: closely related African-origin value systems that emphasize the interconnectedness of individuals with their surrounding societal and physical worlds. ● Techno-solutionism: Lens that sees problems as fixable through technological innovation and a market-centric approach to social change. ● Inclusive Design: Framework based on the simple principle that designing for the widest range of people creates better designs and benefits everyone.

- Technology in Conflict: Mindset that explores how technology can foster conflicts or how technology could improve conflict resolution.
- Subtle Persuasion: Lens that explores how technology, like AI, can alter the beliefs of their users
- Decolonization: Approach that centers on reframing world views from a non-Westernized approach
- Accountability/ Transparency: Framework that explores how to foster clarity, openness, and ownership around processes.
- Critical
- Human Centered Design: Lens that places the user at the heart of the design process.
- Collective Crisis Intelligence: Framework that combines methods that gather information from communities affected by crises and frontline responders using artificial intelligence (AI) for more effective crisis mitigation, response and recovery.
- Participatory Modeling: Framework that focuses on exploring the implicit and explicit knowledge of stakeholders to create formalized and shared representation(s) of reality.
- Human-in-the-loop: Framework that focuses on optimizing models and algorithms through human intervention and contribution.
- Participatory Data Stewardship: Lens that explores the responsible use, collection and management of data in a participatory and rights-preserving way.
- Participatory Action Research: Lens that prioritizes the value of experiential knowledge for tackling problems caused by unequal and harmful social systems, and for envisioning and implementing alternatives.

Methods	Describes the methods used in the project to ensure participation. We found them by looking closely into the method section or project description. We only added those we found in our projects. Mixed Methods (Quantitative & Qualitative) Not Specified Alliance Creation Qualitative Approach Round Tables Interviews Focus Groups Working Groups / Workshops Surveys Request for Public Consultation Sandboxing Training Agile design and prototyping Role-play simulation Gamification Quantitative Observations Cooperative Inquiry

	Panel / Forum Co-design Critical Analysis Case Study Grassroots Organizing Crowdsourcing Data Capturing / Labeling Content/Document Analysis Literature Review Ethnography Pairwise Ranking Contextual Inquiry Red Teaming
Additional Description/Context	If needed, add an additional description of the project or the methods. Not required.
Funding	Describe the type of funding received for the project, if specified. It includes: Multilateral Organizations (IDRC, CAF, BID) International Development Organizations (Hivos, Avina) International Cooperation (ICRC) Research Centers Public Funding Private Funding Awards, Fellowships, Grants Universities, and Academia Unspecified