From Participants to Co-Researchers: How AI Literacy Builds Community Trust and Power in Participatory AI Design

Keeley Crockett
Department of Computing and Mathematics
Manchester Metropolitan University, Manchester, UK
ORCID: 0000-0003-1941-6201

Annabel Latham
Department of Computing and Mathematics
Manchester Metropolitan University, Manchester, UK
ORCID: 0000-0002-8410-7950

Rochelle Taylor
Department of Computing and Mathematics
Manchester Metropolitan University. Manchester, UK
ORCID: 0000-0002-3675-1530

Kaleem Mohammed
Department of Computing and Mathematics
Manchester Metropolitan University. Manchester, UK
ORCID: 0009-0004-7174-4147

***Abstract*—This paper explores the use of participatory AI as a mechanism for increasing public trust in AI by involving community members in the co-creation and co-production of AI projects. It challenges the traditional top-down approach to AI development, arguing that public confidence and trust is directly influenced by the opportunity for citizens to actively shape and understand the technologies that impact their lives. We present a synthesis of research and case studies illustrating the efficacy of this approach, specifically in community-based initiatives. The results show that AI literacy enables informed participation, which in turn strengthens individual agency, thereby supporting greater trust and confidence in AI systems, Together, these dynamics highlight the critical role of AI literacy in fostering a more equitable and socially responsible future for artificial intelligence.**

Keywords—Participatory AI, AI Trust, AI Confidence, AI literacy

I. Introduction

Trust in Artificial Intelligence (AI) refers to the confidence that an AI system will consistently act in a fair, transparent, and beneficial manner, respecting individual privacy and societal values without causing unintended harm. AI and Trust continue to be a heavily researched interdisciplinary topic [1][2][3] spanning theoretical practices from social science, empirical studies and attempts at measurement. In 2025, comprehensive work on AI Governance [4] summarized that trust was not simply a property of a model, people wanted credible evaluation, independent audits, redress mechanisms, and clear lines of accountability underpinned by AI legalisation.

Different user stakeholders will have varying expectations and levels of trust in the use of AI, depending on several factors including their lived experiences, personal knowledge, context of use and understanding of technology. Additionally, they may be unaware of their rights regarding protection from AI and digital technologies or may be excluded from these conversations altogether due to digital poverty. The challenge is how to build trust with the most vulnerable in society, including traditionally marginalized communities where the barriers to engagement require more resources than organisations can or will allocate. A critical concern is that individuals and communities with limited digital access or low levels of AI literacy risk being excluded from meaningful participation in shaping AI's trajectory. Such exclusion not only perpetuates existing inequalities but also undermines democratic principles by silencing diverse perspectives in governance and policy discourse. Ensuring that all members of society can engage with AI—both as beneficiaries and as contributors to decision-making requires deliberate strategies to overcome barriers related to digital infrastructure, education, and socio-economic status. Participatory AI [5] involves engaging diverse stakeholders throughout the AI lifecycle (whole or part) through co-creation and shared decision-making to ensure alignment with societal values and prevent exclusion. Providing opportunities for citizens to act as co-researchers, participate as advisers, consultants and/or designers on an AI project empowers individuals, giving them greater confidence in their own ability to question the use of this technology in everyday lives. Despite the need for a base level of data and AI literacy in order to participate, this should not be seen as a barrier, but rather it should be integrated as part of citizen onboarding as the first stage of the participatory process.

This paper presents three distinct case studies which adopt participatory AI approaches with stakeholders including public sector digital inclusion teams, local councils, and people from marginalized communities. Each case study represents a separate real-world project, applying different participatory methods. The commonality is that they all include citizens from traditionally marginalized communities [6]. Qualitative and quantitative evaluations are presented for each project which demonstrate that participatory engagement can increase citizen trust and/or confidence in AI and digital technologies. Through the case studies, this paper seeks to address the following research question - *In what ways does increased knowledge of data-driven technologies and AI literacy among community participants influence their ability to act as co-researchers in AI design and development processes?*

The three case studies provide empirical and qualitative evidence that participatory AI methods increase public trust and confidence in AI technologies. We demonstrate that the involvement of marginalized communities in the co-creation of AI projects leads to a measurable increase in citizen's confidence and trust in technology and improves citizens ability to use generative AI in a safe and responsible way. Our work challenges the traditional top-down approach to AI development, arguing that participatory AI is a critical mechanism for ensuring the voices and lived experiences of vulnerable communities are represented in AI policy and design. We show that increased AI literacy enables communities to move beyond being passive users of technology to active co-researchers who are empowered to question and shape the digital systems in their lives. By synthesising real-world projects, the paper offers a scalable model for participatory AI governance that prioritises social responsible and genuine community inclusion. However, we acknowledge that within the three case studies the sample is small, but not unexpected in projects of this nature.

This paper is organised as follows. Section II presents related work on AI regulation in relation to the importance of building public trust. In addition, a brief summary of global public perceptions of AI is presented, highlighting the inclusion gap when seeking a representative perspective. Section III outlines the three distinct case studies, the evaluation methodologies and the results in terms of impact on citizens confidence and trust. The paper concludes with a brief analysis and discussion of key findings.

II. Related work

A. AI Regulation and Public Trust

This section provides a brief snapshot of global AI regulation and how it emphasizes public trust as a core component. In regulation, public trust in AI involves adherence to a set of core ethical principles which can cover accountability, fairness, safety, security and robustness, transparency and explainability, protection of human rights and contestability and redress [7,8]. The EU AI Act aims to build public trust in AI technologies though its risk-based regulatory framework [9]. The Act (*Recital 27)* explicitly endorses the 2019 Ethics Guidelines developed by the High-Level Expert Group on AI, which highlight seven core principles for trustworthy AI: human agency & oversight, technical robustness, privacy & data governance, transparency, diversity/fairness, societal/ environmental well-being, and accountability [10]. These principles are foundational to the Act's normative structure, embedding human trust at its core. A core objective of the UK Government AI regulation: a pro-innovation approach relies on five non-statutory principles (safety, transparency, fairness, accountability, contestability) with a specific emphasis on increasing public trust in AI [7]. Experts warn these are “aspirational” and lack legal obligations or sanctions, creating little incentive for compliance [11]. High scale deployment, such as the increased application in public service without sufficient guardrails, means that failures in critical areas (e.g., policing, healthcare) would severely damage trust and delegitimize public institutions.

At the end of 2025, there was no single comprehensive federal AI law, but over 50 US states had introduced over 100 AI-related bills with public trust in AI high on the agenda. For example, California’s SB 53: Transparency in Frontier AI Act, explicitly states the law is designed to “build public trust” through requiring transparency from frontier AI developers. [12]. SB 53 also requires “public disclosure of safety frameworks, incident reporting, and whistleblower protections” [12] ensuring accountability and boosting credibility. However, an Executive Order on "Ensuring a National Policy Framework for Artificial Intelligence" (December 11, 2025) [13] was enacted to establish a single federal AI policy to prevent a "patchwork" of state laws that could hinder innovation. It is too early to know whether centralization through a consistent, national approach to compliance with offering to reassure the public that AI systems are governed uniformly or may damage public trust by suggesting economic advancement is prioritized over safety and fairness. South Korea’s Framework Act on the Development of Artificial Intelligence and Establishment of Trust (Law No. 20676) will come into force on 22nd 2026). The Act sets out a basic plan which includes “*measures for establishing a framework of trust ensuring the fairness, transparency, accountability, and safety of AI*” [14]. Article 29 emphasizes creating a foundation of trust for the safe use of AI through education of safe and trustworthy AI society. Japan has currently a light touch policy foundation of trust for the safe use of AI through the Act on Promotion of Research and Development, and Utilization of Artificial Intelligence-related Technology [15]. The Act aims to appropriately mitigate various risks inherent to AI and ensure citizens’ safety and trust.

Whilst public trust is emphasized as a key criteria of legislation, the challenge is around how to a) operationalize methods and approaches to build trust with citizens, and b) ensure that hard to reach individuals, such as those in marginalized communities, those in digital poverty, or those digitally excluded are not disadvantaged in accessing the benefits of utilizing AI in every-day life.

B. The Nature of AI Trust

Globally, public perception and trust in AI has been captured in annual public surveys. Table I presents highlights of public surveys conducted in 2025 on AI, where n = number of participants. For each survey a brief finding relevant to AI trust is quoted. Such surveys reveal a complex interplay of acceptance, scepticism, and conditional trust. While there is broad willingness to engage with AI technologies, there are concerns regarding reliability, fairness, and ethical governance. The surveys in Table I generally reveal that trust in AI is not uniform: it varies significantly across geographic, demographic, and experiential dimensions. Respondents in advanced economies exhibit greater skepticism compared to those in emerging economies. Regulatory frameworks emerge as a critical enabler of trust, with strong public support for laws and oversight mechanisms. Age-based differences in expectations are evident: younger people tend to emphasize corporate accountability, whereas older individuals place greater confidence in regulatory bodies. Barriers to adoption are dominated by doubts about the integrity of AI-generated content, alongside concerns about data privacy and ethical standards. Generational divides further shape engagement, with older adults demonstrating lower interest in using AI technologies. Experience with AI appears to mitigate perceived risks, as frequent users report lower levels of societal concern compared to non-users. Institutional trust remains fragile, particularly in relation to government use of AI. Interestingly in [21], perceptions of fairness favour algorithmic systems over human decision-making.

Table I Trust in AI Surveys

Survey/ Region	*n*	Snapshot Public Trust Findings
KPGM, Trust, attitudes and use of AI [16], 47 countries	*n=48000*	"54% are wary about trusting AI. While most people feel both optimistic and worried about AI, 72% accept its use. People in advanced economies are less trusting (39% vs. 57%) and accepting (65% vs. 84%) compared to emerging economies" [16]
Ada Lovelace, Alan Turing, How do People feel about AI [17], UK	*n=3,513*	"The majority of the public (72%) indicate that laws and regulations would increase their comfort with AI technologies….. while younger people (18–44) favor company responsibility, those over 55 prefer regulators, reflecting differing levels of trust and expectations based on age." [17]
Tony Blair Institute for Global Change – what the UK thinks about AI [18], UK	*n= 3,727*	"38 per cent of respondents cite a lack of trust in AI content as a barrier, making it the biggest single obstacle to AI adoption. Concerns around data privacy and ethical standards are other leading barriers across all age groups, whereas disinterest in using the technology is a distinct issue among those aged 65 or older….. Only 26 per cent of weekly AI users view AI as a societal risk, compared to 56 per cent of non-users." [18]
Nesta's Centre for Collective Intelligence [19], UK	*n=137*	"Less than half (40%) of the UK public trust the public sector to use AI responsibly" [19]
Ipsos AI Monitor 2025: A 30-country Global Advisor Survey [20]	*n= 23,216*	"Fifty-four percent trust AI not to discriminate or show bias, but only 45% trust their fellow humans to be so fair." [20]
UN 2025 Global Survey on AI and Human Development [21]	*n=21000*	"Public trust in government use of AI varies distinctly by level of human development. In very high Human Development Index (HDI) countries, skepticism prevails, with over half of respondents expressing doubt. By contrast, confidence dominates in high HDI nations, with about two-thirds of respondents optimistic." [21]

The question arises: who are the people who participate in these surveys and are they truly representative of global citizens? Whereas a goal of each survey is to use robust sampling methods to ensure coverage, only [17] and [18] include a mechanism to obtain viewpoints of people who are digitally excluded. The number of respondents for marginalized communities is therefore low and consequently their opinions are excluded. This "inclusion gap" is a significant risk because the five non-statutory government principles outlined in [7] remain aspirational, lacking the necessary legal protections to ensure that the public is protected while opinions of digitally excluded people remain important in the public sector where AI is being used to deliver more efficient and personalized public services. One way to reach marginalized communities is through utilising participatory AI methods. Participatory AI is an effective way to bridge the gap with marginalized communities because it actively involves them in the design, development, and decision-making processes of AI systems, ensuring their voices, lived experiences, and concerns are represented. In this paper, we present three case studies to show how participatory AI functions as this necessary mechanism, providing credible evaluation and accountability that both the public and governance experts demand for the development of trustworthy and socially responsible artificial intelligence.

III. Case Studies using Participatory AI

This section presents three community-based projects that have embedded participatory AI methods. In each case *n* refers to the number of participants. In case studies 1 and 2 community members were compensated for participation with an hourly rate and travel expenses which is a key ethical feature of participatory AI that recognizes participation as a form of labor that demands fair compensation [22].

A. Case Study 1: PEAs in PODs

1) Context

PEAs in Pods is a project [23] which aims to embed public engagement and co-production in AI and data science research. The project trained 23 Public Engagement Ambassadors (PEAs) through a three-pod model: 1) a training programme to upskill data scientists and AI researchers; 2) a mentoring programme focusing on supporting researchers to embed co-production into their own research, and 3) the use of co-creation/ co-production to create three inspirational programs of public engagement in AI marginalized communities.

2) Participants

Communities recruited to work on projects were Back on Track [24], The Tatton [25] and Inspire [26]. Outcomes include a People's Charter for AI [27] *(n=7)*, an educational video (Community Perspectives on Artificial Intelligence) [28] *(n=12),* and practical tools addressing AI risks for vulnerable groups *(n=4)* led by Noisy Cricket [29].

3) Co-creation and Co-design Methods

All activities with communities utilized participatory AI practice through the co-creation and co-production methods. This included onboarding sessions in each community to ensure the project began with a shared language of understanding about what AI is and what it isn't. Recent stories in the media that featured the use of data and AI in either a positive or negative light were used to explore impacts on society with humans. These discussions naturally led to an increase in AI literacy within the co-production sessions, with community participants gaining more confidence to contribute and challenge ideas. The Peoples Charter for AI [27] was co-created over 4 sessions and featured research on public charters, charter principal design exploring what citizens wanted organisations to practically do to conform to each principle, and design of a logo. This was achieved through scaffolded sessions, crafting principles as a team, and creative design work for the logo.

The co-production of the educational video [28] involved community members being active co-researchers rather than interviewees. The community members decided on the responsible AI content and the types of applications that should be included which were appropriate to their communities. Community members also took part in story boarding the video, designing scenes, acting and postproduction reviewing.

The third project with Noisy Cricket explored with community members - *how might we mitigate the risk and strengthen the rewards of using machine learning to help people impacted by homelessness find unconsidered roles and industries based on their strengths?* Together the project team and community members explored AI risk and biases in employment, and, through open discussion, a Socio-Ecological Model was used to capture all perspectives. The interactive sessions focused on how we could help employers use machine learning in recruitment to better understand people's potential, looking at strengths rather than just skills. The outcome was employer training which highlighted how impacted people's data can be used in transparent, trustworthy and practical ways. This was co-designed with Purple Bonsai, user research and learning design studio leading to a practical Strengths Exchange Workshop [30].

4) Evaluation Methodology and Results

In the three PEAs in PODs projects, qualitative evaluation methodologies were used over a period of 2 years to obtain feedback from community participants as they engaged in co-creation and co-production with the public engagement ambassadors (PEAs) and the project team. The methodology involved the use of reflective circles at the end of each co-creation/ production session which were audio recorded and then thematically encoded. As each project was co-produced, evaluation methods did vary. For example, at the Peoples Charter for AI project, community members (*n=7*) took part in a survey at the project onboarding session which explored their individual awareness, attitudes, and experiences related to AI, with specific emphasis on the use of Apps, health (National Health Service) and energy (smart meters). One participant chose not to complete the survey. The survey results for this project are shown in Figure 1.

Surveys were used at the close of all three projects which measured development and perceived confidence of the participants. Community members (*n=21*) engaged with the project took part in a survey in July 2024, where 100% of respondents agreed or strongly agreed that they would like to be involved in projects and opportunities like this in the future and 100% strongly agreed that being part of the project increased their understanding of, and exposure to, universities in Manchester. At the close of the project, one-to-one closing interviews captured deeper insights that the surveys may have missed. The interviews yielded responses that demonstrated how individual community members' opinions and behaviours around AI has changed over the course of the project as evidenced by six different community member quotes:

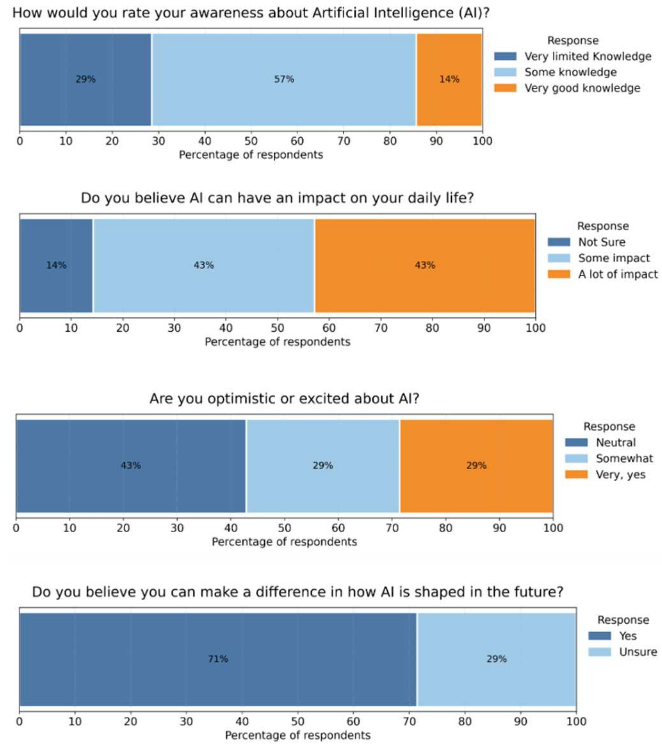

Figure 1 Onboarding baseline survey

- *'I believe that AI is good. For a lot of things. But not for all things. I think it's got its place in society. But not in all areas, for example schools. I had a smartphone for years, and I know about computers and I use internet banking. But I'm worried that we didn't have a say about the AI, its being forced on us. It's the older guys I'm talking about who don't want an AI bot to look after their banking. So being involved in the project has really increased by understanding of where AI is used. I would rather learn, so I can tell people, inform people, so they can feel confident using it. I've gotten other people involved in my community.'*
- '*AI can provide a lot of solutions for broader society, but we need to be asking: Does this solution provide an advantage or disadvantage for ordinary people?*'
- '*I'm a bit skeptical of the high-end AI, where there is no transparency and no way to reverse engineer what has been created. You can't work out what the algorithm is doing. So, if something goes wrong, who fixes it? Are you skewing the information you are putting into it? Is there bias? The people at the top basically seem entitled to do whatever they want because they are wealth creators for the people investing in them. AI development just seems unfettered. We need to rein it in, we need scrutiny and control, and there isn't enough critical thinking.*'
- *'We are living in an AI world, either knowingly or unknowingly. Some of the developments feel unnecessary. AI can provide a lot of solutions for broader society, but we need to be asking the question: Does the solution provide an advantage or disadvantage for ordinary people?.'*

- *'I have more of an awareness of it now. I can see how I algorithms work, I can see: AI has been used here, and how I get nudged towards certain things.'*
- *'Taking part in the project has made me understand a bit more about how AI works. I think its more that I don't trust people (and how they might use AI) then I don't trust the AI itself.'*

Analysis of the qualitative data from the closing interviews demonstrates a growing awareness and confidence in using digital technology and AI literacy amongst community participants. Initial perceptions of uncertainty and concern for rapid growth in AI technologies, which were expressed at project onboarding sessions, had metamorphosed into a greater understanding of AI and feeling more empowered to share their growing knowledge within their community. This reflects how participation in the project has fostered digital confidence, AI literacy and community members to see themselves as active and equal partners in research.

B. Case Study 2: People's Panel for AI

1) Context

The People's Panel for Artificial Intelligence (PPfAI) was designed to bring together a group of citizens to educate, upskill, learn about, discuss and give their views around the use of data and artificial intelligence in people's everyday lives [6, 30]. PPfAI debates the positive and negative consequences of new research ideas, products and/or services pitched to the panel by private businesses, public sector organisations and researchers and provides opinions and feedback from a wider perspective. The Panel is established to ensure that the voice of citizens in communities is heard and to empower them to have confidence in questioning the use of technology [6, 31]. Funded originally by The Alan Turing Institute, members of the public (especially those in traditional marginalized communities) engaged voluntarily in AI literacy roadshows and were given the choice to opt to join the PPfAI two-day training programme in responsible AI to become trained panel members. The participatory nature of this case study lies in the engagement and embedding of citizen voice at multiple stages of the AI-lifecycle through participation in consequence scanning and harms modelling [31]. Panel reports provide an in-depth analysis, plus suggestions for harm mitigations which the organisation is asked to address. Feedback was requested and reported 3-4 months after the panel thus ensuring the loop was closed and the panel was not seen as a tick box exercise for public consultation.

The Manchester City Council's People's Panel for AI, developed in partnership with Manchester Metropolitan University, engages residents—particularly from digitally excluded and marginalized communities—in shaping AI-powered services. In May–June 2024, the Council ran five free, community-based roadshows, attended by over 40 residents aged 18–64+, 17% of whom disclosed a disability. These workshops provided introductory and ethical AI education, resulting in a 30%+ increase in participants' AI understanding, trust in systems, and confidence to critique automated decisions—leading over half of participants to join the Panel [31]. Key lessons included the importance of partnerships with community organisations, targeting outreach via the Digital Exclusion Risk Index, and addressing barriers to access (transport, language, physical accessibility). The Panel now contributes resident insights to AI service design and informs internal policy development. Plans for 2026 include extending participation to VCS and SMEs and embedding challenge-driven public engagement in the city's AI strategy [31].

2) Participants

Residents (*n=67*) from around Greater Manchester and Salford took part in roadshows in a wide range of venues (*n = 6*) including libraries (*n=18*), the Wai Yin Society [32] (*n=10*), Cheetham Hill Housing [33] (*n=2*), No. 93 Wellbeing Centre [34] (*n=11*).

3) Methodology

The methodological approach underpinning the PPfAI was structured as a multi-stage participatory engagement process designed to recruit, train, and empower citizens to critically evaluate AI-enabled products and research. The methodology consisted of four sequential components: community roadshows, panel member recruitment, structured training, and facilitated panel sessions.

The process began with a series of AI Community Roadshows delivered in targeted locations to raise public awareness and provide foundational education on artificial intelligence. The roadshows introduced key concepts—including the role of AI in everyday life, examples of current applications, and case studies for debate—and formally presented the purpose and structure of the PPfAI. These sessions were designed to engage residents from diverse backgrounds and encourage participation from communities that are typically represented.

Participants expressing interest during the roadshows or through community advertising were invited to enroll in two dedicated PPfAI training days. Recruitment materials were disseminated through community centers and public venues to ensure broad accessibility. All individuals who completed the training were eligible to volunteer as panel members, enabling a diverse and inclusive pool of contributors to the final panel discussions. The PPfAI training programme consisted of two full-day interactive workshops that combined instructional content with practical exercises. Training Day 1 introduced participants to the fundamentals of data, machine-learning model development, ethical AI (including bias, fairness, accountability, and explainability), legal frameworks, and the Consequence Scanning technique. Training Day 2 provided applied practice in Consequence Scanning, Harms Modelling, and participation in a mock panel session. Each training day concluded with reflective evaluation to assess participant confidence, understanding, and preparedness for panel participation. Trained volunteers subsequently participated in formal panel sessions lasting approximately two hours. Each session followed a structured format: project pitches by organisations, a facilitated Q&A, independent panel deliberation using the Consequence Scanning framework, and a feedback exchange between panelists and presenters. An independent facilitator documented the discussion and synthesized written feedback summarizing identified consequences, questions, and proposed mitigations. Presenters were invited to respond to recommendations and report back on subsequent changes to their practices or designs. This iterative process aimed to enhance AI accountability and support citizen influence in technological development.

4) Evaluation Methodology and Results

Trust and confidence surrounding AI was measured as part of a larger pre and post set of questions which were conducted at roadshows and training days. Figure 2 shows the pre-and-post scores for two specific questions from the roadshows (*n=67*) and from publics who signed up and completed the two training days (*n=28*). Trust and confidence increased on average by 35% from the roadshows which demonstrates the value that public engagement in local communities around AI literacy can bring to regular citizens. A key feature of the roadshows was that story telling around recent AI related stories in the media gave people the opportunity to debate and discuss ethical principles such as bias and accountability, how it impacted themselves, friends and family, and whether this was a good application of AI ("the machine being in control").

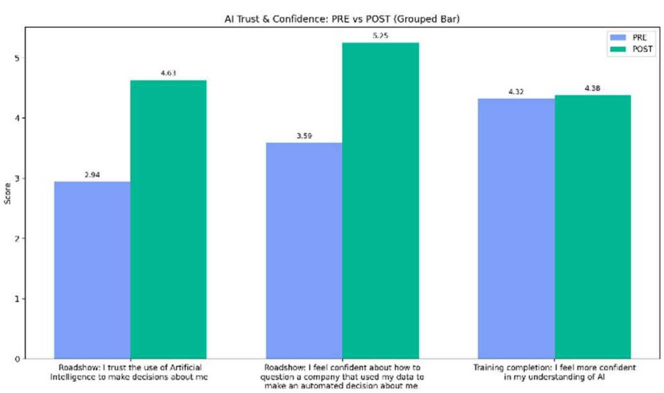

Figure 2 Trust and confidence of publics: road shows and training.

Panellists' confidence continued to grow after taking part in the People's Panels, as illustrated by the following quote:

"*I felt so empowered when the business presenter was writing down what I had said, that he was listening.*"

C. Case Study 3: AI Skills: Train the Trainers Context

1) Context

In this case study, a series of 3-hour workshops was co-created with Manchester City Council (MCC) Digital Strategy Team and engaged charities and social enterprises in exploring practical applications of AI to enhance service delivery and community impact. MCC is a public service organisation supporting over 600,000 residents. The workshops [35] introduced voluntary and community sector organisations to the risks, opportunities, and ethical considerations of AI, everyday use cases, with a focus on Generative AI tools.

2) Participants

Workshops were advertised through Manchester City Councils Digital Inclusion Team to Voluntary, Community, and Social Enterprise organisations (VCSEs) throughout the region in 2 communities. Two workshops (*n=30* and *n=17*) have taken place and were fully booked.

3) Methodology

Participatory mapping and design thinking was used to co-create the content of the workshop that would be suitable for VCSEs with several iterations taking place. The final workshop enabled participants to examine potential benefits and risks, while identifying opportunities for AI to augment existing support mechanisms within local organisations. Each workshop aimed to build confidence, improve digital literacy, and foster informed decision-making around responsible AI adoption in the voluntary and community sector. The session highlighted potential benefits for VCSEs, such as improving operational efficiency, creating tailored support tools, and empowering communities, while addressing limitations and safeguards. "*By creating opportunities to build our collective understanding of AI, we can support more people to be included in our digital society and work collaboratively to close the digital divide in Manchester.*" [36]

4) Evaluation Methodology and Results

Evaluation was conducted by external partners Manchester City Council Digital Strategy Team. Participants answered both a pre- and post-paper-based survey comprising of 5-point Likert and open text questions. Questions 1 to 4 were the same in both pre and post surveys (Table 2) to measure change in knowledge before and after the workshop.

Table 2 Evaluation questions

PRE Questions Likert Questions	
1	I understand what Artificial Intelligence means.
2	I feel confident about how to use generative AI applications such as Copilot and ChatGPT.
3	I am aware how generative AI applications such as Copilot and ChatGPT use my data.
4	I understand that generative AI may not give me correct answers to my questions, and I need to check it.
Post Questions Likert Questions	
1	I understand that generative AI may not give me correct answers to my questions, and I need to check it.
2	I feel confident about how to use generative AI applications such as Copilot and ChatGPT.
3	I am aware how generative AI applications such as Copilot and ChatGPT use my data.
4	I understand that generative AI may not give me correct answers to my questions, and I need to check it.
5	Please write one thing you will do differently as a result of attending this workshop: [text answer]
6	Please estimate how many people and in what capacity you will be able to provide help to as a result of this workshop [text box]

At the first workshop, 20 organisations engaged with 30 participants (*n=30*). Comparing the responses to the first 4 questions: 66% of respondents reported a greater understanding of what AI means after the workshop (Q1); 22% of respondents reported a big increase in confidence with AI, moving up three categories (Q2); 81.5% reported an increased awareness of how

AI uses their data (Q3); understanding of Generative AI may not give correct answers 23 rated strongly agree after workshop, 11 rated strongly agree before workshop (Q4). In terms of reach and impact beyond the workshop (Q6), the number of people VCSE's intend to support is shown in Figure 3.

Figure 3: Reach of Train the Trainer Workshop 1

In capturing behaviour change (Q5), participants stated that the workshop highlighted the importance of not sharing personal data in a GenAI system, being more aware of privacy settings and be more aware of how such systems could benefit day to day work and life tasks.

IV. Discussion and Conclusion

Addressing digital exclusion and fostering AI literacy are not peripheral concerns; they are foundational to creating an inclusive and ethically grounded AI ecosystem. By amplifying community voices and fostering ethical, inclusive AI practices, the three case studies demonstrate a scalable model for participatory AI governance and socially responsible innovation. Each case study utilized different participatory AI methods and different periods of engagement, from one-off road shows and workshops to engagement in 9-month projects. Evaluation pre and post engagement has shown that regardless of method, trust and/or confidence surrounding the use of AI amongst participants increased. Critically to all case studies were a) a mutual language of understanding of AI and the project/ task b) onboarding sessions which featured basic AI literacy upskilling c) the creation of safe spaces where project teams and community members had equal voice. All methodologies are also scalable beyond the region.

To answer our initial research question, increased AI literacy among community participants successfully enables them to act as co-researchers in the broad AI ecosystem. The limitations of using participatory AI lie with researcher upskilling in co-design, co-production and public engagement and the resources (time and funding) to ensure mutual benefit to both community members and team members. Our key findings can be summarised as:

1. Establishing trust with communities is a resource- and time-intensive process. The time and funding required to build these relationships is more than most organizations are willing to be able to invest, but there must be a mutual benefit to both the community participants and research team and/or organizations.
2. The success of participatory AI lies with researcher upskilling. As evidenced by Case Study 1 (PEAs in PODs), researchers often need additional training to learn how to embed co-production in their work.
3. Broad public engagement methods such as roadshows are effective for raising public awareness but converting that initial interest to long-term commitment in a project is challenging, as demonstrated by the small uptake of people beyond the initial educational sessions in the PPfAI.
4. While global surveys of people's perceptions of AI (Table 1) provide large-scale snapshots of public attitudes toward AI across regions and demographic groups—they inherently operate at a distance from the communities they intend to represent. Surveys typically capture surface-level sentiment: levels of trust, awareness, concern, or acceptance. They cannot easily probe why people hold these views, how lived experience shapes them, or what contextual factors influence perceptions of risk and opportunity. Participatory AI, by contrast, offers depth. Rather than surveying people about AI from a distance, participatory approaches engage community members directly in exploratory dialogue, collaborative design, iterative testing, and shared decision-making. This shifts their role from passive respondents to co-researchers who shape the AI systems that will ultimately affect their lives. Depth emerges not only from richer data, but from a transformation in agency: participants articulate concerns in their own vocabulary, interrogate value trade-offs, reveal context-specific risks, and identify design failures invisible to formal questionnaires.
5. Global surveys frequently fail to represent digitally excluded groups who are often the people who will be most impacted by public sector AI. Participatory AI can fix the inclusion gap but struggles with large-scale data needed for traditional policymaking.
6. Even when projects are specifically designed for marginalized groups, several barriers still hinder participation. Lessons learned from the PPfAI include the ongoing need to address obstacles such as transport, language barriers, and physical accessibility. Without deliberate strategies to address these concerns, participatory models still risk excluding the most vulnerable.
7. The need for additional time for onboarding community members to ensure that there is a base level of AI and digital literacy is something that participatory models need to consider but this adds complexity and time to projects, which must be factored in.

V. Conclusions and Further Work

The movement to mainstream participatory AI practice reflects a growing recognition that traditional, top-down approaches to AI development are insufficient for ensuring fairness, equity, and public trust. AI systems need to be designed and shaped with communities rather than just imposed on them. Efforts to democratize AI emphasize that genuine public participation is essential to uphold fairness and accountability [37]. For example, community-led AI auditing in areas such as criminal justice and public service delivery [38]. To build a participatory AI ecosystem new organisational practices need to embed participation as part of the culture of responsible AI, ensuring barriers to engagement are lowered and resourced. Additionally, there needs to be a commitment to equity, ensuring digitally excluded and marginalized groups are actively involved rather than silently omitted from data and design processes. Community members who took part in the projects described in this paper continue to participate in further research projects. This includes the co-creation of a Responsible AI charter for small businesses, advising around practical AI Accountability in different contexts and giving their opinions on UaI (Understandable AI) interfaces. Furthermore, they have become AI ambassadors in their own communities. The demand for community participation is there, as researchers we just have to bridge the gap.

Acknowledgments

PEAs in PODs were funded UKRI Engineering and Physical Sciences Research Council (EP/W033488/1). The Peoples Panel for AI was funded by The Turing and Manchester City Council. We would like to thank all our community co-researchers who took part in our projects.

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), pp.1-30.

[2] Benk, M., Kerstan, S., von Wangenheim, F. and Ferrario, A., 2025. Twenty-four years of empirical research on trust in AI: a bibliometric review of trends, overlooked issues, and future directions. *AI & society*, *40*(4), pp.2083-2106.

[3] Lacity, M.C., Schuetz, S.W., Kuai, L. and Steelman, Z.R., 2025. It's a matter of trust: Literature reviews and analyses of human trust in information technology. *Journal of Information Technology*, *40*(3), pp.286-313.

[4] ITU, The Annual AI Governance Report 2025: Steering the Future of AI., Available: https://www.itu.int/epublications/publication/the-annual-ai-governance-report-2025-steering-the-future-of-ai

[5] Delgado, A., & Callén, B. (2023). *Participatory AI: Engaging publics in the co-production of algorithmic systems.* AI & Society, 38(1), 45–59. https://doi.org/10.1007/s00146-022-01394-4

[6] Latham, L. Crockett, K. Towards Trustworthy AI: Raising awareness in marginalized communities, *2024 International Joint Conference on Neural Networks (IJCNN)*, Yokohama, Japan, 2024, pp. 1-8, doi: 10.1109/IJCNN60899.2024.10650957.

[7] UK. Gov, AI regulation: a pro-innovation approach, 2023, [online], Available: https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach

[8] OECD, Principals on Trustworthy AI, 2023, Available: https://www.oecd.org/en/topics/sub-issues/ai-principles.html

[9] EU AI Act, 2024, Available: https://artificialintelligenceact.eu/

[10] Ethics Guidelines for Trustworthy AI, (2019), Eurpoeam Commission, Available:https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html

[11] The AI regulation gap: Is the UK's pro-innovation approach enough?, (2025), [online], Available: https://www.computerweekly.com/news/366636561/The-AI-regulation-gap-Is-the-UKs-pro-innovation-approach-enough

[12] Governor Newsom signs SB 53, advancing Californias worl-lerading artificial intelligence indutry (2025), Available: https://www.gov.ca.gov/2025/09/29/governor-newsom-signs-sb-53-advancing-californias-world-leading-artificial-intelligence-industry/

[13] White House (2025), ExecutiveOrder: ENSURING A NATIONAL POLICY FRAMEWORK FOR ARTIFICIAL INTELLIGENCE, Available:https://www.whitehouse.gov/presidential-actions/2025/12/eliminating-state-law-obstruction-of-national-artificial-intelligence-policy/

[14] South Korean Ministry of Government Leglidsation, Framework Act on the Development of Artificial Intelligence and Establishment of Trust (2025) Available: https://cset.georgetown.edu/publication/south-korea-ai-law-2025/

[15] Act on Promotion of Research and Development, and Utilization of AI-related Technology" (referred to as "AI Act", (2025), Available: https://www.gov-online.go.jp/en/assets/hj_november_2025_p24-25.pdf.

[16] KPGM, Trust, attitudes and use of AI (2025), Available: https://assets.kpmg.com/content/dam/kpmgsites/xx/pdf/2025/05/trust-attitudes-and-use-of-ai-global-report.pdf

[17] Roshni Modhvadia, Tvesha Sippy, Octavia Field Reid and Helen Margetts, 'How Do People Feel About AI?' (Ada Lovelace Institute and The Alan Turing Institute, 2025) https://attitudestoai.uk/

[18] Mökander, J. Puddick, J. Goel, N. Wager, A. Rhydderch, T. Margetts, H. Cameron, D. Roff, B., What the UK Thinks About AI: Building Public Trust to Accelerate Adoption, 2025, Available: https://institute.global/insights/tech-and-digitalisation/what-the-uk-thinks-about-ai-building-public-trust-to-accelerate-adoption

[19] Nesta's Centre for Collective Intelligence (CCI), (2025), Available: https://www.nesta.org.uk/press-release/less-than-half-of-the-uk-public-trust-the-public-sector-to-use-ai-responsibly-survey-finds/

[20] The IPSOS AI Monitor 2025, Available: https://www.ipsos.com/sites/default/files/ct/publication/documents/2025-06/Ipsos-AI-Monitor-2025.pdf

[21] UN 2025 Global Survey on AI and Human Development, Available: https://hdr.undp.org/data-center/2025-global-survey-ai-and-human-development

[22] Berditchevskaia, A., Peach, K., and Malliaraki, E., 2021. Participatory AI for humanitarian innovation: a briefing paper. London: Nesta. Available: https://www.nesta.org.uk/report/participatory-ai-humanitarian-innovation-briefing-paper/

[23] PEAs in Pods, 2026, Available: https://peasinpods.mmu.ac.uk/

[24] Back on Track, 2026, Available: https://www.backontrackmanchester.org.uk/

[25] The Tatton, 2026, Available: https://thetatton.co.uk/

[26] Levenshulme Inspire, 2026, Available: https://www.lev-inspire.org.uk/

[27] Crockett, K., Taylor, R., Henry, J., Thorpe, M., Ngozi, N., Kirby-Hawkins, H., Gowda, S. and Stevens, F., 2024. The Peoples Charter for Artificial Intelligence, 2024, Peoples Charter for AI, Avaiable: https://doi.org/10.23634/MMU.00637617

[28] Crockett, K., Thorpe, M., Henry, J., Taylor, R. and Ngozi, N., 2024. PEAs in Pods: Community Perspectives on Artificial Intelligence, Available: https://e-space.mmu.ac.uk/637664/.

[29] Noisy Cricket, 2026, Available: https://noisycricket.org.uk/

[30] Purple Bonsai, (2026), Available: https://www.purplebonsai.co.uk/

[31] Crockett, K. Latham, L. Wood, M. Abberley, L. Rawsthorne M. Attwood, S. Building Trust – The People's Panel for AI, *2023 IEEE Conference on Artificial Intelligence (CAI)*, Santa Clara, CA, USA, 2023, pp. 168-170, doi: 10.1109/CAI54212.2023.00080.

[32] Wai Yin Society, (2026), Available: https://www.waiyin.org.uk/

[33] Cheetham Hill Housing (2026), Available: https://www.nmcp.org.uk/contact-us/

[34] No. 93 Wellbeing Centre, (2026), Available: https://www.gmmh.nhs.uk/no-93/

[35] Manchester Digital Strategy, Train the Trainers, Available: https://www.linkedin.com/feed/update/urn:li:activity:7323325510019710976/?utm_content=&utm_medium=email&utm_name=&utm_source=govdelivery&utm_term=

[36] Local Government Association, Manchester City Council: People's Panel for AI: Building skills, confidence and residents' voice, 2025, Available: https://www.local.gov.uk/case-studies/manchester-city-council-peoples-panel-ai-building-skills-confidence-and-residents

[37] Taylor, R.R., Murphy, J.W., Hoston, W.T. et al. Democratizing AI in public administration: improving equity through maximum feasible participation. AI & Soc 40, 3653–3662 (2025). https://doi.org/10.1007/s00146-024-02120-w

Hassan, S., Asad, S. M. N., Eslami, M., Mattei, N., Culotta, A., & Zimmerman, J. (2024). PACE: Participatory AI for Community Engagement. Proceedings of the AAAI Conference on Human Computation and Crowdsourcing, 12(1), 151-154. https://doi.org/10.1609/hcomp.v12i1.31610