Please cite the Published Version

Additional Information: This is an accepted manuscript of a conference paper accepted for presentation at IEEE World Congress on Computational Intelligence 2024

PEAs in PODs: Co-Production of Community Based Public Engagement for Data and AI Research

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

Edwin Colyer
Scientia Scripta
UK
edwin@scientiascripta.co.uk

Lauren Coulman
Noisy Cricket
UK
lauren@noisycricket.org.uk

Caitlin Nunn
Manchester Centre for Youth Studies
Manchester Metropolitan University, UK
ORCID: 0000-0003-3145-309

Sarah Linn
Manchester Centre for Youth Studies, Manchester Metropolitan Univeristy, UK
ORCD: 0000-0003-2974-0152

Abstract— The operationalization of ethical principles, current and emerging legalization, and the understanding and mitigation of potential consequences to individuals and society are key challenges in the design, development, and deployment of Artificial Intelligence (AI) driven systems. As part of forthcoming global legislation, organisations, businesses, and researchers developing AI solutions as a service or innovating novel applications will need to openly address ethical principles such as bias, fairness, explainability, transparency, data privacy, accountability, and safety through AI Governance. Developing responsible AI is essential to building public trust, yet few AI researchers know how to engage and capture the insight of diverse people, especially those with lived experiences. This paper presents, an operationalization model for Co-production and Public Engagement in the field AI which focuses on Voice, Value and Variety to guide how academics and businesses might approach meaningful public engagement in the context of research and development of AI products and services. The model is applied within a project called PEAs in PODs which seeks to empower the research and development community to engage meaningfully with traditionally marginalized communities on the subject of AI. Recommendations are made for the integration of co-production methods into AI research to enable meaningful public engagement for all AI researchers.

Keywords— Participatory AI, Co-production, Public Engagement, Trustworthy AI, Human Centered AI

I. INTRODUCTION

The application of AI (Artificial Intelligence) in people's everyday lives continues to gain momentum, against a backdrop of emerging AI legalisation and clear directives stemming from global ethical guidelines and principles that are essential to build "Citizen trust" [1,2]. Trust itself comprises of a set of complex human behaviours which can be defined as an emergent property, built from social interactions between two or more agents, one who trusts and the other who is trusted [3]. Adding AI to the trust equation has led to the development of a series of frameworks by organisations and governments on what constitutes Trustworthy AI [3,4], i.e. The EU's high-level expert group stated that trustworthy AI has three components: 1) it should be lawful 2) it should be ethical and 3) it should be robust [4]. Kaul et al, [4] identifies that there is still a significant implementation gap of such frameworks and the transferability from research to practice. Different user stakeholders will have different expectations and trust levels of the use of AI in specific domains and applications, based on several factors including their lived experiences (for example, denial of a service → high street banks close → services provided through apps/online banking → excludes those with low digital literacy), level of personal knowledge, skills and understanding of what AI actually is. In addition, they may not know what their human rights [6] are with regards to protection against AI and digital technology or may be excluded from such conversations altogether due to digital poverty. As an innovator/provider of AI products and services, to truly build trust, the involvement of people (beyond the conceptual target user group), specifically from marginalised communities is essential. To achieve this goal, we need to adapt techniques such as co-creation and co-production into the AI ecosystem and embed public engagement.

Participatory AI can be defined as "the involvement of a wider range of stakeholders than just technology developers in the creation of an AI system, model, tool or application." [7]. Nesta [7] found that case studies of participatory AI tended to be academically driven and focused on stakeholder engagement in limited elements of the AI lifecycle. Typically, the four levels of participatory AI are a) Consultation; b) Contribution; c) Collaboration; and d) Co-creation/co-design [7]. Within collaboration [8] cites co-production, citizen science and public engagement as mechanisms to utilise shared expertise in problem solving. [7] illustrates the levels of participatory AI through a series of real-world case studies including human in loop machine learning, but there is a lack of detail or examples with regards to the operationalization of participatory AI such as recruitment of participants, onboarding (safeguarding and informed consent) and continual evaluation and reflection. The Ada Lovelace Institute provides three examples of public participation in action [9] in a report that focuses on participatory AI within commercial AI labs to ensure tech is

beneficial to society, but the detail on how to operationalise this is missing. Birhane et, al. [10] argues that if AI is participatory then it must be acknowledged that members of the public, external to those stakeholders engaged within the AI lifecycle, have knowledge, skills and motivation to actively engage. Delgado et al [11] reported that there were few empirical studies of what good participation in the AI process looked like in terms of robust domain specific design. Whilst Zhang et al. [12] highlights the challenges of engaging different types of stakeholders in terms of guidelines, power dynamics and equity awareness in the context of AI in decision making.

This paper provides insights and learnings of a project, known as PEAs in Pods [13], that uses co-production of community based public engagement for AI research to demonstrate that community engagement can make research more pertinent and impactful. The project seeks to empower the research community (universities and businesses) to engage meaningfully with traditionally marginalised communities and embed coproduction methods into individual and organisational research processes and governance. Traditionally marginalised communities include geographic socioeconomic deprivation, older people, who are often less engaged and under consulted regarding the impact of technology, people with lived experience of homelessness and ethnically diverse communities. Engagement and empowerment within the project is being achieved through the delivery of public engagement activities around the theme of ethical Artificial Intelligence and data-driven technologies. The paper presents and applies (in the context of the PEAs in PODs project) an operationalization model for Co-production and Public Engagement in the field AI to guide how academics and businesses might approach meaningful public engagement. A key challenge of this project is in establishing a mutual language of understanding of AI between members of the public and university researchers. Furthermore, there is strong influence on public perception based on how AI is portrayed by the media which can act as a baseline for highlighting concerns and raising awareness. A second challenge is in how to communicate, in non-technical language ethical AI principles (e.g., bias, accountability, explainability etc), current and emerging AI regulation and impact on human rights to explore in depth what people's concerns are. The key contributions of this paper are:

- An operationalization model for Co-production and Public engagement in ethical and responsible AI.
- Initial findings on how bespoke training on community engaged research equips AI and data science Early Career Researchers (ECRs) with appropriate tools and knowledge to work effectively with communities.
- Recommendations for the integration of co-production methods into AI research to enable meaningful public engagement.

This paper is organised as follows: Section II provides an overview of the methods of participatory AI, co-production, and public engagement; Section III to VI provides an overview of the PEAS in pods project and methodology, details of POD activities, including training of academic researchers and small businesses with initial evaluations. Finally, Section VII provides a series of recommendations in how to integrate co-production and public engagement into AI research.

II. BACKGROUND AND RELATED WORK

A. Participatory AI

The key challenges in participatory AI can be summarized as 1) a lack of understanding from stakeholders of what it is and where it can be applied within the AI lifecycle (which can be application domain specific); 2) the barriers to public participation and who the public themselves should be; There are few participatory AI frameworks that exist. Nesta [7] published a framework for operationalizing participatory AI in 2021 as part of its humanitarian innovation briefing paper which highlighted, three case studies of participatory AI in practice. The framework comprised of 4 key stages of opportunity for public participation within the AI lifecycle: Collaborative problem framing, data collection and curation, algorithm development and deployment and oversight. These opportunities have been developed from examination of good practices from a small number of case studies. Subsequently, this led to four levels of participatory design of AI:

- Consultation – participation occurs outside the core AI development process (e.g., consequence scanning) which may have little impact on the design.
- Contribution – Selected stakeholders' complete necessary independent tasks (i.e., data set creation, data labelling, evaluation) contributing towards AI development but have no interaction with each other or the development team.
- Collaboration – several opportunities to participate within the AI lifecycle with interaction with technical team (i.e., evaluating model fairness, AI performance)
- Co-creation – participation take part in the entire project lifecycle and have an over watch role.

[7] concludes that participatory AI must allow for people to have "meaningful engagement" and be part of a bigger picture in development of responsible and trustworthy AI including addressing digital literacy challenges and ensuring privacy and safety of the public participants.

Research undertaken by Ada Lovelace considered where public participation might occur within the traditional AI lifecycle starting with problem formulation through to monitoring and evaluation of an AI system postproduction [9]. The work also emphasized the lack of shared terminology around what participatory AI actually is and what methods could be employed and concludes "*While participation is not a silver bullet, it is an important tool to ensuring data and AI debate and practice is in step with people's attitudes, opinions and concerns about technology.*"

The importance of citizen trust in the use of AI has been strongly highlighted in emerging legislation [14.16] yet there has been no concrete guidance on how this can be achieved within our diverse global society. Participatory AI approaches provide opportunities for researchers and businesses to embed public concerns, worries and societal harms from initial

ideation of new products and services. For example [28] drew attention to the benefits of an SME engaging in consequence scanning in the early stages of product design with a community of older people which led to AI product design changes.

B. Co-production

The original concept of co-production stemmed from the development of public services where co-production was the process of working together to find a shared solution to a problem [17]. Effective co-production is built on:

- Power and privilege – both in building relationships and sharing power within the team [18]. Clearly defined roles and responsibilities are needed from the outset.
- Ownership – mutually understanding that the research, product and service which is co-produced is jointly owned.
- Diversity and inclusiveness of people in bringing together different lived experiences, perspectives, and skills on the co-production journey.
- Respect for everyone's voice to be heard equally.
- Mutual value for everyone in taking part and in benefitting from the project outcomes.

Lepratte and Yoguel [19] showcased the co-production process as part of the overall technological solution for digital innovation in health services. While there is some overlap between co-production methods and those within participatory AI, there are also distinct differences. Applying co-production within the field of AI requires researchers, members of the public, communities, and practitioners to meaningfully engage throughout the whole AI project lifecycle where it is essential that all stakeholders can communicate using a mutual language of understanding of what ethical and responsible AI is, are aware of their own power and positionality in forming working relationships and share power and responsibility in achieving the project goals.

C. Public Engagement in AI

"*Public engagement describes the myriad of ways in which the activity and benefits of higher education and research can be shared with the public. Engagement is by definition a two-way process, involving interaction and listening, with the goal of generating mutual benefit*", National Co-ordinating Centre for Public Engagement (NCCPE) [20]. General frameworks for public engagement exist, for example, the NCCPE Framework has a focus on research conducted in universities linked to research impact [21]. It is important to identify the benefactors of public engagement [22], where the value of engagement in AI should be clear to members of the public and not just a tick-box exercise by a company associated with its values. [23] recommends narrative building as a way to facilitate public participation and engagement in public sector services that use Algorithms, data, and AI, however there are no examples of where the method has been applied. Groves [9] conducted research exploring public participation in commercial AI labs and recommended that commercial businesses should engage in partnerships with community groups. Park [24] focuses on the ethical engagement of AI, recommending that engagement should 1) Allocate time and resources to promote representation 2) Adopt inclusive strategies prior to AI development 3) Ensure

training is provided on the integration of ethics. A further deep systematic search of both academic and grey literature (beyond the space of this paper) indicates that whilst there are frameworks and methods of what good public engagement looks like in the field of AI, there are extremely limited published examples and methods.

III. OVERVIEW PEAs IN PODs

A. Aims and high-level objectives

The PEAs in PODs project has four aims: 1) to increase the public engagement and coproduction skills and confidence of Public Engagement Ambassadors (PEA)s through training, reflective mentoring and "learning-by-doing"; 2) is to increase knowledge about data-driven technology and AI research among community participants and hence create the conditions for community members to participate as active stakeholders in research and design processes.; 3) demonstrate the benefits of coproduction methods to the R&D community, as a powerful way to align research to ethical principles and real-world societal needs, especially those of traditionally marginalised communities and 4) to build sustained relationships between PEAs, research institutions and traditionally marginalised communities and embed such interactions into institutional research processes.

A high-level overview of the project objectives can be found in Table I. Our ambition is for this project to lay the foundations for significant long-term impact across the tech research community as a model for ethical, effective coproduction of AI solutions. The project started in September 2022 and runs for 33 months. NB: Each POD comprises of a number of PEAs.

Table I High-level Objectives PEAs in PODs

Obj	Brief Description
1	Train 20 researchers from universities and small-medium businesses in co-production through the development of a novel interdisciplinary training programme on public engagement skills, coproduction, and evaluation, within the context of AI research (POD 1).
2	Deliver and support PEAs with quarterly 1h reflective mentoring sessions. Our principle is to "leave no PEA behind", so all PEAs involved in Pod 1 training will be invited to participate in these sessions, even if they do not participate in Pod 3 (POD 2).
3	Plan, design and deliver a participatory ideation, matching and pitching event with PEAs and community participants to co-create ideas for three separate, coproduced, community based programmes of inspirational public engagement events/activities about data science and AI research (POD 3).
4	Development of 12-month public engagement action plans with each PEA to encourage on-going involvement between the researchers and community groups or others (POD 2).
5	Guide PEAs and community co-researchers to coproduce, deliver and evaluate a programme of inspirational public engagement activities (POD 3). Teams of PEAs worked with community co-researchers on one of three community driven projects. They will be required to commit a minimum of 20 hours development and delivery time each to their projects. Each community project will have a budget of £7k.
6	Use evaluation of POD 3 to coproduce a "Data Ethics and AI in a Box" legacy resource for on-going self-guided learning and researcher engagement about data-driven and AI technologies and ethics across communities.
7	Co-create with the host and partner universities, PEAs and communities', mechanisms for an Institutional Change agenda to

integrate sustained researcher-community interaction into research processes, initially in AI and Data Science Research.

The scope of this paper will focus on contributions relating to POD 1 and POD 3.

IV. POD 1: RECRUITMENT AND TRAININIG

A. Recruitment Roadshows

Recruitment for the project was conducted between December 2022 – January 2023 through seven information 'roadshows' (three UK Universities) and flyers targeting PhD students and Early Career Researchers (Research Assistants and Lecturers) working either within the broad field of AI or data driven technologies. The roadshows attracted 60 Eventbrite signups and 5 additional people attended in person. The aim of the roadshows was firstly to introduce the PEAs in Pods project, to define and provide the benefits of public engagement and co-production with respect to Global and National AI strategies [1], and ethical AI principles [2] and to pitch the opportunity to become a Public Engagement Ambassador. 26 roadshow attendees completed the roadshow evaluation surveys which comprised of six Likert style questions and two open text questions. 90% of roadshow attendees said the roadshow had encouraged them to undertake future public engagement activities to communicate their research and 90% of roadshow attendees said that the roadshow had encouraged them to apply to be a PEA in the project and to pursue collaborative work within their own research. An open text question asked attendees "What in your opinion, is the most exciting thing about the PEAS in PODS Project?". Answers included:

- "Engaging with people - accessing their needs - hopefully to design responsible tech"
- "To have a real conversation with the citizens regarding AI. Gaining the public engagement skill".
- "Considering approaches to engage in public collaboration around AI".
- "To learn about co-production in AI and get to practice it".

Interested attendees applied to become a PEA through completion of an application form comprising of 15 questions. A key criterion was that applicants had to be conducting research working in AI and/or data driven technologies. Table II shows a selection of responses to the question – "How do you see communities benefitting in public engagement and collaboration around the theme of artificial intelligence?" demonstrating that the applicants understood that the value of public engagement in AI must also be felt by the communities which are engaged.

Table II How can communities' benefit from PE? (Selected responses)

I think communities will benefit greatly from public engagement and collaboration around the theme of artificial intelligence. Misinformation and hype are widespread with respect to artificial intelligence and often direct attention away from the real issues it creates (E.g., loss of privacy, environmental impact, loss of dignity). Furthermore, global challenges such as climate change and COVID-19 have made the need for effective communication with respect to science and technology clear. Through public engagement and collaboration, members of wider society will be better able
to understand, question, and make informed decisions with respect to artificial intelligence.
As AI makes inroads into ever more aspects of daily living it is important that people understand what AI can do and what it cannot do as well as what it shouldn't do.
Many people don't know the impact of AI in their daily life. Even though their data massively using and exploited in many ways, many are unaware of the effect of AI and data-driven technologies. Some of the publics are unaware of the positive impacts of AI, they always see AI from a distance and have a fear to use them. As a PEA, we can help to use AI and new technologies positively and ethically for a better society.
Awareness will help them make informed decisions.
Bridging the digital divide and enable the communities to enjoy the immense potential, opportunities and capacity as offered by AI.
As AI becomes prevalent and an integral part of all systems that the public will use, the community by voicing their apprehensions, and understanding etc. will ensure that they can play a role in shaping dialogues to frame policies that are responsible, fair, and with minimal bias for the broader society.
Communities - specifically some of the communities that I have typically worked with - often feel marginalised, alienated, or irrelevant to conversations about AI, but I have really enjoyed and valued encouraging reflections and opening up discussions in focus groups, talks, and various other channels. One of my concerns with AI and robotics is the notion of sleepwalking into a technologically determined future that people fear, so community work and public engagement not only ignites interests and passions but also does some management of the misconceptions and dystopian fears (and indeed utopian hopes) around these technologies that we often encounter.
By actively engaging with communities, researchers can better understand their work's impact and gain insight into the diverse perspectives and needs of the community. This will enable more responsible use of AI resources that serve the need of the community. Furthermore, involving communities in the research process will increase public trust and support for the ethical use of AI. Lastly, it will lead to more creative and impactful research outcomes.
It will be easier for communities to engage and participate in what they can understand, and this is where the program comes into play. Now people will know what their data can offer in terms of suggestions that can improve the quality of their lives and the efficiency of their workflows. In addition, they will be able to gain a deeper understanding of artificial intelligence and identify it easily within the community. Moreover, I believe that this will encourage people to participate in research studies and contribute where they are able.
Developing community networks of shared understanding and support networks for knowledge transfer. Develop community confidence to work together to find solutions and make AI accessible and less daunting for marginalized communities.

B. Bespoke Training

From the 22 recruited candidate PEAs, 18 were able to fully commit and undertake 4 x 4-hour training modules over June and July 2023, with additional sessions on Power and Positionality, Onboarding and Evaluation. The training sessions were developed by the interdisciplinary (social science, impact, AI engineers) project team including Noisy Cricket [25] and Scientific Scripta [26], emphasising the value of PE and public engagement to all stakeholders (PEAs, communities, publics, and academic institutions. In addition, an independent community member consultant with specific lived experiences engaged with all training sessions. Table III lists the aims of each session. Full details of the training programme will be in provided in a forthcoming publication.
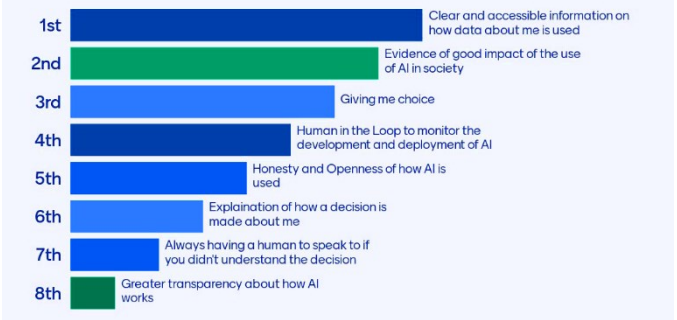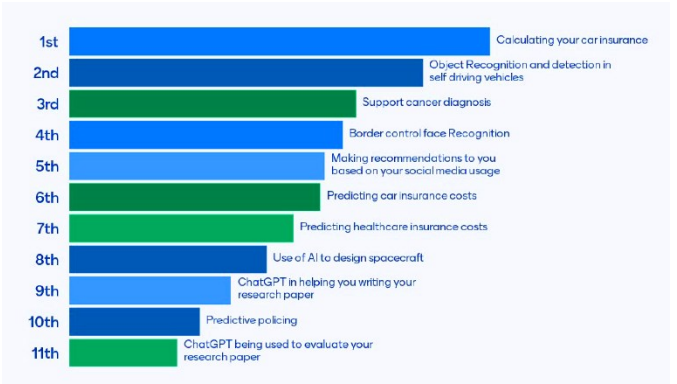
The sessions incorporated both social science and computer science pedagogy, adopting an active learning approach through a performance lens [27]. This was essential as a key objective for the project was to ensure that PEAs could adapt elements of their learnings into their own research. Sessions embedded the

ethical AI (Data driven technology) theme and were very interactive allowing confidence in ability to grow through practice with performance opportunities coming in POD3 (Section VI).

Table III Overview of PEAs Training Programme

Session 1	Introduction to collaborative working and community engagement in the context of ethical AI.
Aims	• To co-create a shared ethos and understandings for working together and start building collaborative relationships • To explore and establish shared understandings of ethics and legal frameworks in AI research. • To introduce community engagement and collaboration approaches, including history and theories • To evaluate PEAs knowledge and understanding of ethical AI practice • To model and provide opportunities to practice ethos, tools, and practices through session delivery.
Session 2	What is community?
Aims	• To explore different understandings and types of community • To explore the ethics and practice of engaging and collaborating with communities • To explore community engagement and co-production in the establishing phases of a project • To evaluate PEAs knowledge and understanding of community • To model and provide opportunities to practice ethos, tools, and practices through session delivery.
Session 3	Public engagement (PE) and co-production in practice
Aims	• To continue exploring approaches to working with communities • To explore community engagement and co-production in conducting research • To explore institutional barriers and consider systems change. • To evaluate PEAs knowledge and understanding of PE • To model and provide opportunities to practice ethos, tools, and practices through session delivery.
Session 4	Public engagement and co-production in research dissemination and impact
Aims	• To explore community engagement and co-production in research dissemination and impact • To explore what 'ending well' looks like • To explore community engagement and co-production in systems change • To evaluate PEAs knowledge and understanding and the training process • To model and provide opportunities to practice ethos, tools, and practices through session delivery.

For example, in training session 1, one component was an interactive session exploring ethical and trustworthy AI. Participants were asked three general questions on their perception of trust: Q1: Did they trust some applications of AI rather than others; Q2: How would they rank the main themes in terms of importance for trustworthy AI; Q3: What three words would they use to describe the barriers to Trustworthy AI. Results are shown in Figures 1 to 3 respectively.

Obtaining car insurance premiums was seen as the most trusted application amongst PEAs, whilst the use of ChatGPT in both writing and evaluating research papers and the use of predictive policing had lowest trust (Fig.1). Figure 2 shows that the most important consideration in whether they conceived the use of AI to be trustworthy concern the use of personal data with the AI lifecycle whilst there was less concern about transparency. The word cloud in figure 3, shows that the main barrier to trusting the use of AI was understanding, followed by unease about bias.

Figure 1: PEAs perceptions of trust of specific AI applications

Figure 2: PEAs perceptions – ranked themes for Trustworthy AI

Figure 3: Barriers to Trustworthy AI.

A further example activity in Training Session 2, required PEAs to draw their data and AI lifecycle as they would practice within their own research and identify where they think publics could participate. This was seen as a challenging exercise, due to the black box nature of machine learning algorithms, but also the difficulty in trying to use plain and simple language to communicate complex ideas to people with no formal training in AI. Evaluation was built into the end of each training session,

not only to enable reflective practice, common within the social sciences but less so within computer science communities, but also to improve the content of the training sessions. From the PEAs perspective, key learnings included.

- The role of intersectionality - a framework for understanding oppression [18], in the context of ethical AI.
- The many dimensions of power and privilege and its distribution when working with communities.
- Examples of research co-design in practice [27] [28].
- Understanding how to engage with different communities through a deeper understanding of what a community is.
- Knowledge gained from each session and access to the training materials that could be applied within own research and context.
- Confidence and empowerment to lead activities during the training sessions such as the use of Ketso kits.
- The importance of having a community of practice initiated through the PEAs training programme to allow PEAS to apply knowledge in real community engagement projects.

From the 18 participants that were trained in community engaged research, 12 committed to working with communities as PEAs (target was 10) in POD3 of the project. This included representation from three universities. The 12 PEAs engaging with POD3 subsequently attended a series of additional workshops exploring power and privilege and onboarding designed and provided by partner Noisy Cricket [25] to prepare them to actively engage with marginalized communities.

V. POD 3 COMMUNITY PROJECTS

The aim of POD 3 was to allow PEAS to experience public engagement and co-production through working with three communities to develop coproduced, community-based programmes comprising of resources/events/activities about data science and AI research. PEAs would work alongside traditionally marginalised communities to explore the impact of AI on/within their communities and to co-produce an output. Each project was also advised by an independent community member consultant with specific lived experiences. The Each project was allocated a £7k budget. The projects commenced in May 2023 and aim to be delivered by July 2024. The communities engaged are:

- Back on Track (Central Manchester)–a charity which provides community learning for disadvantaged adults [29]. The project involves working with people with lived experience of multiple disadvantage' (e.g., homelessness).
- Inspire (Levenshulme, Manchester) – a social enterprise charity which provides community-led services that promote the well-being of people [30]. The project focuses on the older members of the community.
- The Tatton (Odsal, Manchester) – a charity whose aim is to "develop the capacity and skills of the members of the socially and economically disadvantaged community of Odsal (Salford) in such a way that they are better able to identify, and help meet, their needs and to participate more fully in society." [31]

A. POD 3 – Operationalization Model for Co-production and PE in AI

The proposed operationalization model for engaging with communities has been devised from frameworks for participatory AI, theories of co-production adopted from the social sciences and real-work operationalization from social impact consultancy, Noisy Cricket (project partner). At the heart of the model and indeed the ethos of co-production is value of engagement to community members and their representative communities. An overview of the model can be seen in Figure 4. The model reflects that the pathways to community engagement may be different depending on whether a project/challenge or problem that requires community co-production is known in advance; whether there are existing relationships between researchers and community members, or indeed significant pre-work is needed to find communities who would seek to benefit from talking part. Once community members are recruited and onboarded, the pace of engagement, should reflect the speed in which the community wants to progress with the project. For a researcher, this can be quite challenging to balance project timelines with community needs but it is essential to building trust with the community. The model comprises of the 6 key tasks briefly described in Table IV and illustrated using the PEAs in Pods project.

For each hour of engagement, community members were compensated £15 per hour in vouchers. In addition, for two-hour meetings, refreshments were provided and in three-hour workshops lunch was also provided. Travel expenses for participation was also included. It is worth noting that for participants with disabilities, taxis were provided.

B. POD 3 community projects high level overview

PEAs worked in groups with recruited members from each community. The three community projects are currently in the early stages of co-production. This section provides a brief description of each project, but due to the nature of the co-production process, amendments may be made. The Back on Track project is focused on Co-Production for Empowering People Disadvantaged by Homelessness: Mitigating Artificial Intelligence Risks and Strengthening Rewards. Within this project, co-production will be used to first understand the impact of AI used in recruitment through engagement with communities disadvantaged by homelessness to understand their experiences and needs. The project will also explore solutions on how such individuals can be matched with unconsidered job roles and industries based upon life and lived experience skills. The Inspire project is focused on the creation of "The Peoples Charter for AI". Initial co-production sessions have examined existing citizen charters relating to public services, the aims of the proposed charter, impacted stakeholders and how it can incorporate ethical AI principles. The Tatton project involves creation of educational resources (including videos) to raise awareness of AI in society (i.e., AI Assistants such as Siri, Alexa, on-lining banking etc), demystifying AI and help community members feel safe in AI around them in their everyday lives. There is a strong emphasis

on raising awareness of use of data in AI models, data privacy, who owns insight from the model, how consent is given and the requirement for Terms and Conditions on the use of AI in applications to be communicated in simple language and no jargon.
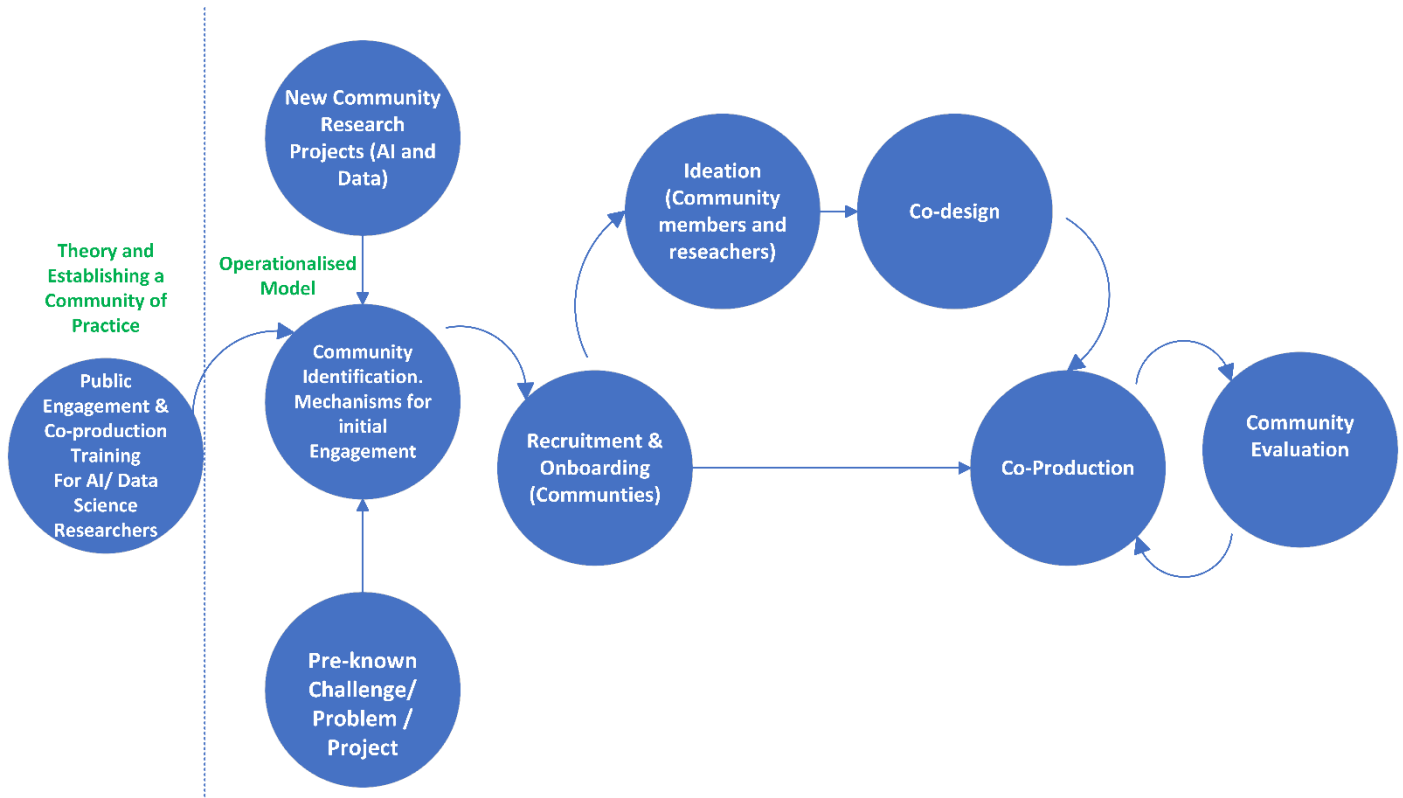

Figure 4: Operationalization Model of PE and Co-Production

Table IV Tasks

Task	Description
1	**Recruitment** of members of the community to take part in one of the 3 community projects. Methods included recruitment surveys, short presentations involving AI stories in the media and conversations in community centers. In this phase there is a fundamental requirement to clearly state the value to people of being involved, compensation offered for participation as a volunteer, clearly define their role and management of expectations.
2	**Onboarding** – recruited community members first attend an onboarding session where they completed informed consent to take part in a project, engaged in icebreaker activities with the PEAs/ research team and then explored ethical AI and data related topics, problems and challenges that mattered to their community. Safeguarding, data privacy and data collection, roles and responsibilities were clarified with plenty of Q&A opportunity.
3	**Co-creation phase** - Depending on the project and community, a series of ideation sessions were required and led by PEAs to clearly establish the problem and challenges that the community wanted to address within their project in relation to ethical AI. *Note*: Back on Track community members had a clearly defined problem centered around how the use of AI was utilized when people impacted by homelessness were searching for jobs. Thus, this project moved from onboarding straight to co-production phase.
4	**Co-production phase** – this phase consisted of, on average, 4 * 2-hour meetings with PEAS and community members to co-produce outputs for community public engagement programs that would address community problems/challenges around the ethical use of AI. Depending on the output(s), these meetings might involve consequence scanning, design, role-play, interviews, questionaries, experimentation, prototyping ideas, and evaluation.
5	**Delivery and wider community evaluation** – Each community project would include at least one launch event within the community, where feedback could be attained from the wider community. Individual project evaluation objectives were co-produced to ensure that communities would benefit from the project which would in turn encourage future engagement.
6	**Sustainability and impact** - Post project launch, this phase would look at how each project and associated outputs could be used to sustain and increase community awareness of ethical AI and give people the confidence to question the use of AI in their everyday lives.

C. POD 3 community projects high level overview

PEAs worked in groups with recruited members from each community. The three community projects are currently in the early stages of co-production. This section provides a brief description of each project, but due to the nature of the co-production process, amendments may be made. The Back on Track project is focused on Co-Production for Empowering People Disadvantaged by Homelessness: Mitigating Artificial Intelligence Risks and Strengthening Rewards. Within this project, co-production will be used to first understand the impact of AI used in recruitment through engagement with communities disadvantaged by homelessness to understand their experiences and needs. The project will also explore solutions on how such individuals can be matched with unconsidered job roles and industries based upon life and lived experience skills. The Inspire project is focused on the creation of "The Peoples Charter for AI". Initial co-production sessions have examined existing citizen charters relating to public services, the aims of the proposed charter, impacted stakeholders and how it can incorporate ethical AI principles. The Tatton project involves creation of educational resources (including videos) to raise awareness of AI in society (i.e., AI Assistants such as Siri, Alexa, on-lining banking etc), demystifying AI and help community members feel safe in AI around them in their everyday lives. There is a strong emphasis on raising awareness of use of data in AI models, data privacy, who owns insight from the model, how consent is given and the requirement for Terms and Conditions on the use of AI in applications to be communicated in simple language and no jargon.

Each of these projects are at different stages. For example, at the time of writing the Inspire and Tatton projects have engaged in the co-production phase, whilst Back on Track have just completed onboarding community members. One key challenge is in establishing a mutual language of understanding of what AI is and what it isn't. PEAs engage every day in complex language related to their research, whilst community members may only come across terms such as models, algorithms, inference from data and machine learning within engagement sessions and it is often a requirement to have to revisit what is meant by AI. Having an operationalized model (Figure 4 and Table III) is good for scaffolding, but individuals and communities will move at a pace which suits them to ensure that they continue to get value from the engagement. Consequently, the PEAs must be adaptive and patient. A major challenge for the research team is having a funded project with strict deadlines/outputs and trying to apply this to a project where the community members hold equal power.

VI. INTEGRATION OF CO-PRODUCTION INTO AI RESEARCH RECOMMENDATIONS

With a strong international emphasis on mechanisms to build public trust in AI, the PEAs in PODs project has adopted a co-production methodology, inspired by participatory AI to community engagement. We define meaningful engagement as:

- **Voice** – empowering the community with skills to feel able to contribute to the discussion around ethical AI, appreciating everyone comes to the conversation with different lived experiences.
- **Value** – ensuring that the time and unique insights invested by community members during co-production sessions meets their needs and requirements and contributes to their wider community.
- **Variety** – recognizing that diverse community members have different perspectives and inclusive co-production can lead to more creative problem-solving and innovation.

To-date the key learnings from applying the operationalized model across three community projects can be summarized in a series of recommendations that will help AI researchers to consider the impact of their work though meaningful engagement. The recommendations are as follows:

- Co-design, co-production training and public engagement training should be a core part of PhD skills training programme for any AI and data driven research that has an impact on people.
- Public Engagement activities need to be seen as real added value embedded within personalised PHd programmes from the start, to enable students to be able to communicate and understand the impact of their work to diverse publics. Critical for research funders too!
- Time (number of sessions/ workshops/meetings plus time for preparation and evaluation) and rescources (compensation, travel costs, room hire, catering) of all stakeholders (the public, researchers, support staff) need to be costed from the outset of any research project. Where possible, consult the communities that will be involved and ask what good looks like for them. Where possible co-production should take in a venue chosen by the community.
- Public engagement and co-production training should be mandatory for supervisory teams, and all academics that have a core responsibility for research. Core training for all with a domain specific emphasis is required. Senior academics should lead by example.
- Understand the barriers and expectations to working with communities. This may be from your own power and positionality, establishing a mutual language of understanding not just about AI, but in relation to the community problem(s)/challenge(s). Seek support from institutional social scientists and reach out in advance to community networks especially to people with lived experiances and listen.
- As a researcher, be prepared to change, adapt and be flexible when it comes to any form of public engagement and be prepared to navigate and challenge institutional processes and procedures.

VII. CONCLUSION AND FURTHER WORK

One of the motivations for the PEAS in PODS project was to build people's confidence and awareness of ethical AI,

specifically, enabling traditional marginalized communities to have a voice in this field. Secondly, we wanted to enhance capacity and confidence of researchers in community engagement (through training, mentoring, and learning by doing). Currently the trained PEAS are embarking on building meaningful relationships with communities through three community projects, the outputs of which will form part of a forthcoming (2025) open source "product" known as AI Ethics in a Box. The journey is starting to nurture an appetite for change – PEAs become advocates and change-makers, communities feel empowered to interact and institutions see the benefits of change. This paper has shown the value and benefit of co-production and PE to early career and PhD researchers engaged in AI research and how they are embracing the challenge of how to build citizen trust through responsible AI. This paper has shown how bespoke training on community engaged research equips AI and data science researchers with appropriate tools and knowledge to work effectively with communities through applying an operationalized model. This work is not for the fainthearted and requires a new skill set for the traditional AI researcher. Further work will involve completion, delivery, and evaluation of the three community projects, publication of our PEAs in PODs training programme and a co-produced publications with community members.

ACKNOWLEDGMENT

REFERENCES

[1] United Nations Educational, Scientific and Cultural Organization, 2021, [onine], Recommendations on the Ethics of Artificial Intelligence, available: Recommendation on the Ethics of Artificial Intelligence - UNESCO Digital Library, Date accessed [20-02-23]

[2] OECD AI Principes Overview (2023), [online], Avaiablle: https://oecd.ai/en/ai-principles; Date accessed: [21-12-2023].

[3] Hancock PA, Kessler TT, Kaplan AD, Stowers K, Brill JC, Billings DR, Schaefer KE, Szalma JL. How and why humans trust: A meta-analysis and elaborated model. Front Psychol. 2023 Mar 27;14:1081086. doi: 10.3389/fpsyg.2023.1081086. PMID: 37051611; PMCID: PMC10083508.

[4] EU High-levek Expert Group on AI , Ethical Guidelines for Trustworthy AI (2019), [online], Available: https://www.aepd.es/sites/default/files/2019-12/ai-ethics-guidelines.pdf, Date accessed: [21-12-2023]

[5] Kaur, D. Uslu, S, Rittichier, K. Durresi, A. 2022. Trustworthy Artificial Intelligence: A Review. ACM Comput. Surv. 55, 2, Article 39 (February 2023), 38 pages. https://doi.org/10.1145/3491209

[6] Equality and Human Rights Commission, AI safeguards 'inadequate', watchdog warns (2023), [online], Available: https://www.equalityhumanrights.com/ai-safeguards-inadequate-watchdog-warns#:~:text=Addressing%20the%20equality%20and%20human,being%20exacerbated%20by%20algorithmic%20biases.

[7] Berditchevskaia, A., Peach, K., and Malliaraki, E. (2021). Participatory AI for humanitarian innovation,London: Nesta. Available: https://media.nesta.org.uk/documents/Nesta_Participatory_AI_for_humanitarian_innovation_Final.pdf

[8] National Co-ordinating Centre for Public Egagement, Quality Practce, (2023), [online], Available https://www.publicengagement.ac.uk/quality-practice

[9] Groves, L. Going Public:Exploring public participation in commercial AI labs, [online], Available: https://www.adalovelaceinstitute.org/report/going-public-participation-ai/

[10] Birhane, A., Isaac, W., Prabhakaran, V., Diaz, 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, pp.1-8.

[11] Delgado, F. Barocas, S. Levy, K. 2022. An Uncommon Task: Participatory Design in Legal AI. Proc. ACM Hum.-Comput. Interact. 6, CSCW1, Article 51, https://doi.org/10.1145/3512898

[12] Zhang, A., Walker, O., Nguyen, K., Dai, J., Chen, A. and Lee, M.K., 2023. Deliberating with AI: Improving Decision-Making for the Future through Participatory AI Design and Stakeholder Deliberation. Proceedings of the ACM on Human-Computer Interaction, 7, pp.1-32.

[13] PEAS in PODs (2024), [online], Available: https://peasinpods.mmu.ac.uk/

[14] EU AI Act, 2023, [online], Available: https://artificialintelligenceact.eu/ Date accessed [20-02-23]

[15] UK Gv, A Pro-innovation approach to AI regulation, (2023), [online], Available: https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper

[16] Candada Gov, Artificial Intelligence and Data Act, (2023), available: https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act

[17] Bandola-Gill, J., Arthur, M., & Leng, R. I. (2023). What is co-production? Conceptualising and understanding co-production of knowledge and policy across different theoretical perspectives. Evidence & Policy, 19(2), 275-298, https://doi.org/10.1332/174426421X16420955772641

[18] Andersen, N., 2022. Diverse Examples and Balanced Perspectives. Enhancing Inclusion, Diversity, Equity and Accessibility (IDEA) in Open Educational Resources (OER)

[19] Lepratte, L. and Yoguel, G., 2023. Artefacts, routines, and co-production: a pioneering case of artificial intelligence-based health services in Argentina. Industry and Innovation, pp.1-23.

[20] NCCPE, High Quality Engagement Framework, (2023), [online], available: High Quality Engagement Framework | NCCPE (publicengagement.ac.uk), Date accessed [04-01-24]

[21] UKRI, Public Engagement strategy 2023, [online], Available https://www.ukri.org/what-we-do/public-engagement/our-strategy/

[22] Wilson, C., 2022. Public engagement and AI: A values analysis of national strategies. Government Information Quarterly, 39(1), p.101652.

[23] Marmolejo-Ramos, F., Workman, T., Walker, C., Lenihan, D., Moulds, S., Correa, J.C., Hanea, A.M. and Sonna, B., 2022. AI-powered narrative building for facilitating public participation and engagement. Discover Artificial Intelligence, 2(1), p.7.

[24] Park, T. (2022) Making AI Inclusive: 4 Guiding Principles for Ethical Engagement (Partnership on AI, 20 July 2022) <https://partnershiponai.org/paper/making-ai-inclusive-4-guiding-principles-for-ethical-engagement/, Date accessed [07-01-24]

[25] Noisy Cricket, (2023), [online] available: Noisy Cricket – Building people powered movements

[26] Scienti Scripta, (2024), [online], Avaiable: Science communications agency - Scientia Scripta - impact first

[27] Raj, Rajendra; Sabin, Mihaela; Impagliazzo, John; Bowers, David; Daniels, Mats; Hermans, Felienne; Kiesler, Natalie; Kumar, Amruth N.; MacKellar, Bonnie; McCauley, Renée; Nabi, Syed Waqar and Oudshoorn, Michael (2021). Professional Competencies in Computing Education. In: Proceedings of the 2021 Working Group Reports on Innovation and Technology in Computer Science Education, ACM pp. 133–161. https://www.backontrackmanchester.org.uk/

[28] People Powered AI, (2023), [online], Available: https://www.turing.ac.uk/people-powered-ai?__cf_chl_tk=DWAS_NPOrk2DLbAmhfDr5bKr02.w9z7_4kL4h_s.nLI-1704125364-0-gaNycGzND3s, Date accessed [01-01-24]

[29] Treescapes (2024), [online], Available: Treescapes: "Voices of the Future" Project (mmu.ac.uk). Date accessed [01-01-24]

[30] Back on Track, (2024), [online], Available: https://www.backontrackmanchester.org.uk/. Date accessed [01-01-24]

[31] Levenshulme Inspire, (2024), [online], Available: https://www.lev-inspire.org.uk/ . Date accessed [01-01-24]

[32] The Tatton, (2020), [online], Available: https://find-and-update.company-information.service.gov.uk/company/04893603/filing-history?page=1