AI FOR SOCIAL GOOD?

Inspirations from Participatory Action Research (PAR) to Critical Data Studies

Principal Investigator:
Dr Andrea Medrado
Senior Lecturer, School of Media and Communication
Communication and Media Research Institute (CAMRI)
University of Westminster
a.medrado@westminster.ac.uk

Co-Investigator:
Dr Pieter Verdegem
Reader, School of Media and Communication
Communication and Media Research Institute (CAMRI)
University of Westminster
p.verdegem@westminster.ac.uk

1. What is AI?

AI, artificial intelligence, is computer programming that learns and adapts from data. The term is used to describe any computer system that is taught to mimic intelligent human behaviours[1]. There is a lot of hype around AI but the fact is that AI is all around us, everyday. It's in our phones, in ride hailing apps, recommendation systems, in smart maps, smart cities, facial recognition, biometric technologies and more. A lot of research has been done on what AI can do to solve problems in areas like medicine, engineering, to tackle global challenges, such as the pandemic and climate change.
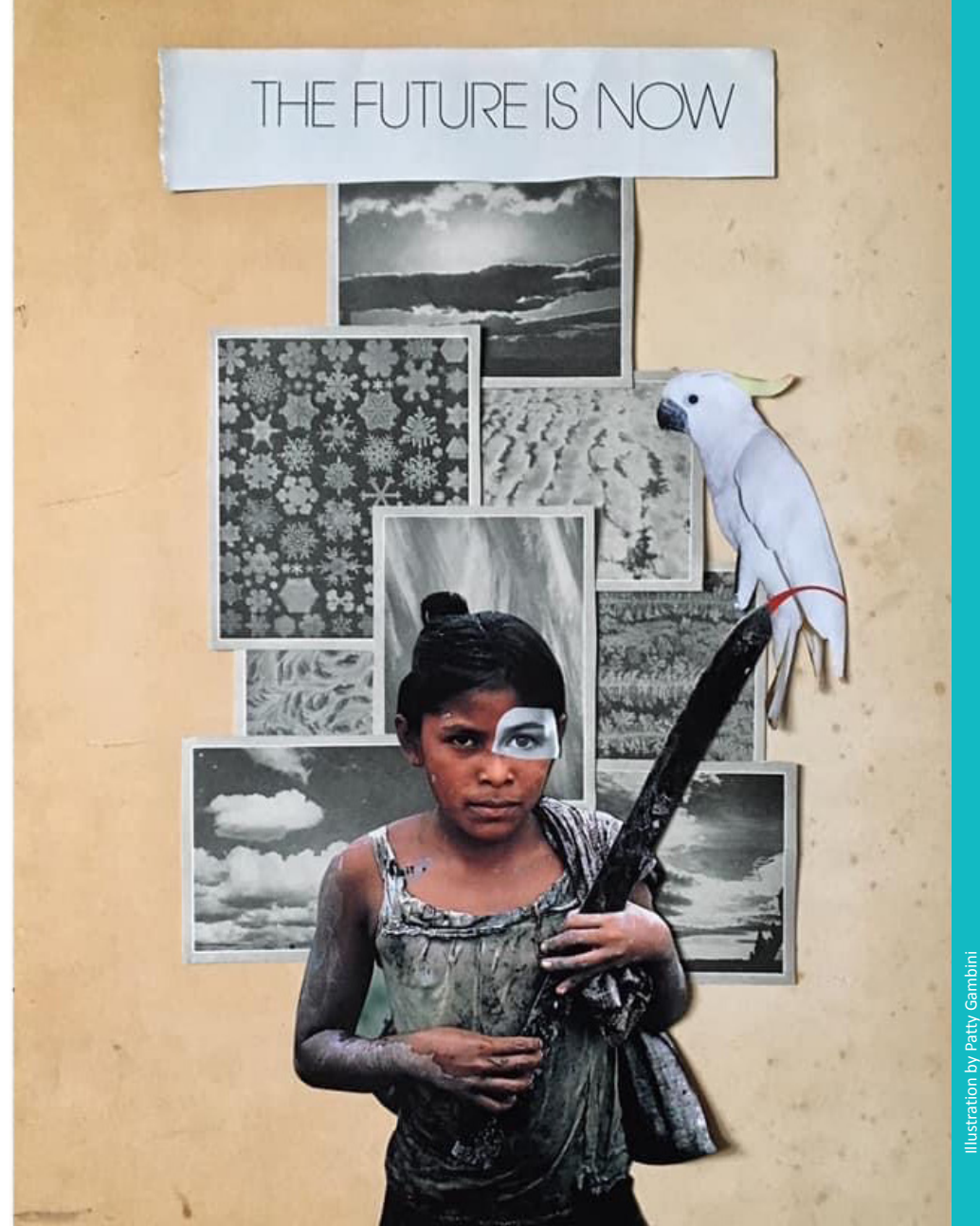
At the same time, there is another story that shows that AI technologies might increase global inequalities, contribute to environmental degradation, and are based on the exploitation of labour and extraction of resources[2]. As such, there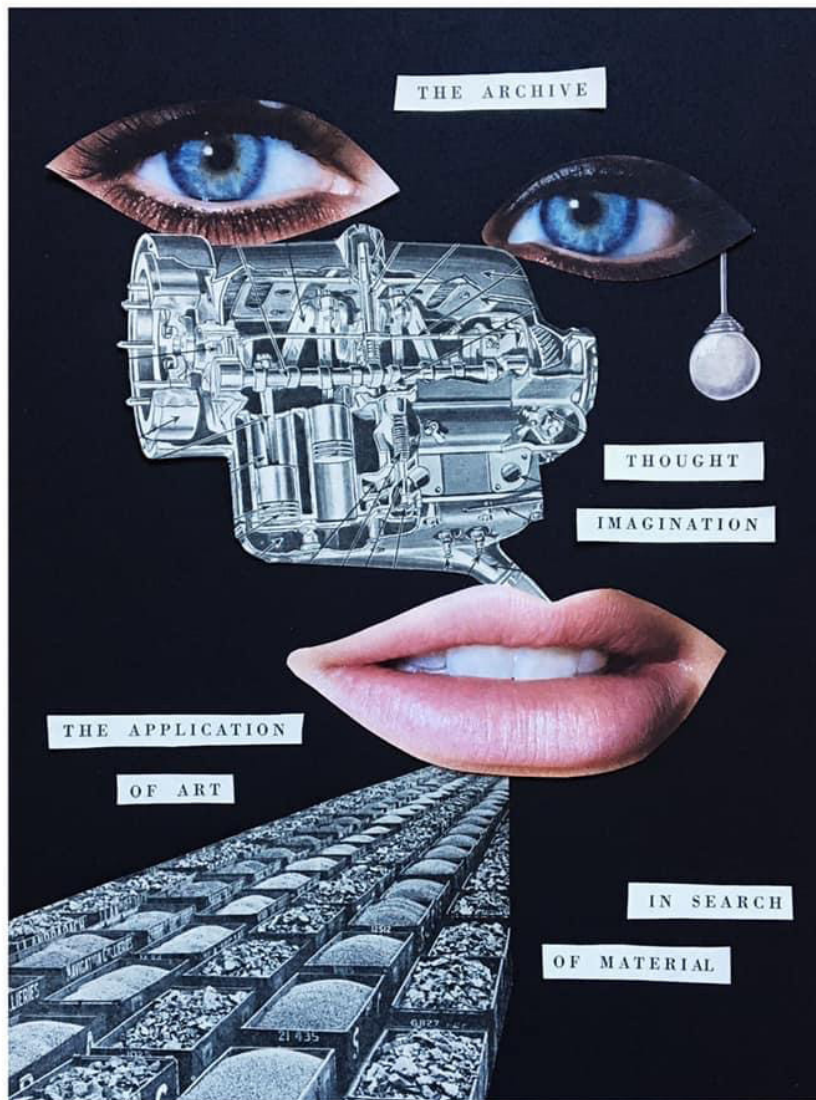 is growing evidence that AI is driving/ intensifying a new colonial order whereby the Global North profits from the exploitation of the Global South[3].

So, what if AI technologies are actually worsening some of the social problems that we face? Problems like discrimination, racism, social inequalities, unemployment, precarious work?

In this project, we reflected upon these and other questions, recognising the need to think critically about:

- **What type of AI do we have? And what type of AI do we want?**

- **How can AI be more inclusive?**

- **Can we speak of "AI for social good" without really knowing what it means to promote social good?**

THE FUTURE IS NOW

Illustration by Patty Gambini

2. AI for Social Good?

We are at the crossroads of technological developments which are changing our economy and society. It is argued that much of our productivity and prosperity will be derived from the systems and machines we are creating. AI is considered to play a central role in this and will thus contribute to offer benefits for society - AI for Social Good.

Corporations in the tech industry often carry out programmes on "tech for good" to contribute to their corporate social responsibility (CSR) goals. Google, for instance, has launched a "Tech for Social Good" initiative, providing funding to non-profit, academic or research institutions, civic entities, and social enterprises in search of technical help and funding for projects focused on sustainability, economic opportunity, or cyber security[4]. The funded projects are expected to address a societal problem and offer solutions to marginalised communities.

However, it is often unclear what social good means. And good for whom? If we think about AI for social good, in simple terms, this can refer to technologies that contribute to societal and human well-being. Aspects such as shared benefits and shared prosperity are key for this[5].

In sum, AI for social good must focus on:

"The design, development and deployment of AI systems in ways that (a) prevent, mitigate or resolve problems adversely affecting human life and/ or the wellbeing of the natural world, and/or (b) enable socially preferable and/or environmentally sustainable developments"[6]. (Floridi et al., 2020).

But to what extent is this definition useful for having a discussion on the ground and in practice about the types of AI that we have and the types of AI we want?

3. Participatory Action Research (PAR)

Our approach was inspired by the Latin American tradition of Participatory Action Research (PAR). This tradition is linked to the work of scholars like Orlando Fals Borda, from Colombia, and Paulo Freire, from Brazil[7]. The development of PAR is linked to social movements in Latin America, such as movements for land reform, anticolonialism, and as responses to developmentalist paradigms in which techno-solutions from the North are tested and applied in top-down ways to the South.

Specifically, this tradition emerges from the struggle of the oppressed, or those who are invisibilised and silenced, to break free from the oppressor (to use terms that are recurrent in Freire's work).

PAR has multiple features but for the purposes of this small project, we focus on three aspects:

1) *Vivência* **or lived experience.** It can be defined as a full experience of an event with all its possibilities. It is lived through direct participation. *Vivência* cannot be observed; it can only be lived, felt, and experienced.

2) Praxis. The general meaning of praxis contains elements of dynamism and change, reshaping ideas into action. Praxis is an act of engaging, exercising, and practising ideas. Praxis also implies a critical reflection, an awareness of the process and its aim. In Freire's pedagogy, praxis refers to the actions taken by the oppressed in the processes of their liberation and paths to freedom.

3) Conscientisation or awareness raising. Oppressed populations reach a tipping point when they are critiquing the actions that they once believed were critical to their survival. This tipping point is the conscientisation or the process in which humans achieve a deep awareness of their socio-cultural realities and of their capacity to transform these realities.

Illustration by Patty Gambini

Why combine Participatory Action Research (PAR) and AI for Social Good?

Many believe that AI is creating a new colonial world order. PAR, as it manifests in the work of Latin American scholars like Fals Borda and Freire, among other scholars from the Global South, represents a decolonial approach to delve into these issues. Often, we use approaches developed in the Global North and apply them to the Global South. In this project, our aim was to investigate how perspectives developed in the Global South can offer alternative ways of thinking about better/fairer AI realities and futures. Additionally, if PAR approaches were originally developed in the 1960s and 1970s, as part of what is considered a radical past, how can they be applied in the present to study AI, which is deemed to be the future?

Here, we also explain what we mean with the term "Global South". The "Global South" does not represent a simple geographic location. Rather, these "South(s)" feature as metaphors for oppression, social inequalities, and human suffering. The term has disadvantages and advantages. On one hand, the "Global South" is problematic. It lumps together the other, leading to more "othering" of the others. By adopting it we might indirectly contribute to homogenising what constitutes the true "South". This is further complicated by the fact that there are Souths in the Norths and vice-versa (such as pockets of poverty in the Global North and pockets of wealth in the Global South). On the other hand, notions of the "Global South" can be useful. The term triggers conversations about colonial legacies, oppression, and marginalisation. In this way, the Global South conveys the meaning of a political solidarity project[. It also expresses an openness to listen to "the other" and to each other as "others". It communicates the willingness to change things for the better, to right what is wrong, to join forces and fight injustices.

4. Workshops "AI for Social Good?": Drawing inspiration from Participatory Action Research:

To answer these questions, in the "AI for Social Good?" Project, we gathered a group of participants – undergraduate and postgraduate students in fields like digital media, journalism, PR, computer science, as well as tech workers and tech activists. We conducted a series of workshops at the University of Westminster, organised as follows:

Day 1 – Stories of AI (Main principles: "vivência", "conscientização").
We shared stories of AI in the everyday life of invisible workers in the Global South. This triggered a conversation about AI, ideas of social good, while we also introduced the PAR tradition to the group. At the end of the day, it was decided that participants would document their own "days in AI" via storytelling, by producing a piece of text, video or collage. Here, we were inspired by projects that apply creative storytelling to studies of technology and marginalised communities[8]

Day 2 – My Day in AI (Main principles: "vivência", "conscientização", empathy)
The participants returned two weeks later having produced a creative piece to document their daily experiences in and with AI. The artefacts were shared among participants so that they could engage with each other's content prior to the workshop. During the session, we asked: which pieces of content produced by others we empathised with and why? This exercise prompted participants to collectively elaborate a list of questions and problems that could be tackled with an action or intervention in workshop 3.

Day 3 – Discussion of AI Intervention (main principles: "conscientização", praxis)
Participants spent more time articulating problems in terms of AI not contributing for social good such as: there are prejudices encoded in AI; AI inputs come from one predominant source (Western, white, male perspective + powerful companies in the Global North); AI-driven recommendation systems interfere with our autonomy; business models do not support dialogue with empathy; AI policy is flawed, we have little information about the big players/actors in AI and what

their agendas are. From these ideas, we selected three main questions to tackle as a group (discussed in the next section).

Day 4 – Design of AI Intervention (main principles: "conscientização", praxis)
Doing collective brainstorming, the participants proposed two creative interventions to tackle the problems with notions of AI for social good. We present them briefly in this report.

5. Questions that stemmed from the workshops:

Inspired by the principle of awareness raising about AI and their own but also different marginalised groups' socio-cultural realities, the participants proposed the following overarching questions:

(1) How can users have more autonomy over their social interactions in a world increasingly impacted by algorithms and automation?

(2) Since notions of empathy cannot be embedded in automated systems, what are the implications for users?

(3) How can we bridge the gaps between users, activists, policy makers and technologists to generate constructive dialogue on AI and data-driven technologies?

6. Initial ideas that stemmed from the workshops

During the "Stories of AI" and "My Day in AI" exercises, participants identified how the mainstream media's reporting and representations of AI are limited. These representations either fall into the hype of AI or tend to offer dystopian Hollywood narrative tropes about AI. As an alternative to this, participants came up with initial ideas to develop a documentary piece which would focus on ordinary users' everyday engagements with AI. The (real) characters in this documentary piece would also engage with other people's recommendation systems, empathising with them and acquiring a greater awareness or consciousness ("*conscientização*") about issues of AI-driven discrimination and exclusion.

The second idea proposed by the project's participants related to issues of media policy. During the workshops, we brainstormed ideas for the development of a platform called "Have you also heard?" which would be embedded in social media platforms like Twitter. The aim was to enable social media platforms to point users to diverse perspectives on the same political issue. The rationale was described as follows: your views are X but have you also heard Y? The group believed that this would help reduce a radicalisation of views which might be caused by algorithmic logics in which users are always exposed to the same types of content.

To conclude, we note that it was beyond the scope of this small project to offer a list of solutions to complex questions such as the ones we raise in this report. Rather, drawing from the PAR principles of *conscientização* (conscience raising), *vivência* (lived experiences) and praxis, the aim was to gather a diverse group of people and think about these issues together. By engaging in these initial dialogues, we also learned about various projects, groups and organisations doing excellent work in issues of AI, ethics, equality and inclusion.

On the next page we offer some suggestions if you wish to find out more.

7. Organisations, projects and groups working with issues of AI, equality and inclusion.

The Alan Turing Institute:

The UK national institute for data science and artificial intelligence

https://www.turing.ac.uk/

Ada Lovelace Institute:

An independent research institute with a mission to ensure data and AI work for people and society

https://www.adalovelaceinstitute.org/

AI4All:

Working to make artificial intelligence more diverse and inclusive, AI4All offers education programmes for underrepresented high school students in partnership with universities like Stanford, UC Berkeley, Carnegie Mellon, and Princeton.

https://ai-4-all.org/

AJL - Algorithmic Justice League:

AJL's mission is to raise awareness about the impacts of AI, equip advocates with empirical research, build the voice and choice of the most impacted communities, and galvanise researchers, policy makers, and industry practitioners to mitigate AI harms and biases. We're building a movement to shift the AI ecosystem towards equitable and accountable AI.

https://www.ajl.org/

Algorithmic Fairness and Opacity Group:

Hosted at the UC Berkeley School of Information, the group conducts research education, develops policy, builds systems, brings theory into practice, and bridge disciplinary boundaries. They centre human values in the design and use of technical systems to support more equitable and just societies.

https://afog.berkeley.edu

Allen Institute for Artificial Intelligence (AI2):

Non-profit organisation with the mission to contribute to humanity through high-impact AI research and engineering. All work is directed towards AI for the Common Good.

https://allenai.org/

Center for AI and Digital Policy:

Nonprofit AI policy & research institute focused to ensure that artificial intelligence and digital policies promote a better society, more fair, more just, and more accountable.

https://www.caidp.org/

Coding Black Females:

The largest community of Black women in tech in the UK. Growing, educating and inspiring the community of Black women in tech.

https://codingblackfemales.com/

Data for Black Lives:

A movement of activists, organisers and scientists committed to the mission of using data to create concrete and measurable change in the lives of Black people.

https://d4bl.org/

Data Justice Lab:

Examines the intricate relationship between datafication and social justice, highlighting the politics and impacts of data-driven processes and big data. The lab is hosted at Cardiff University's School of Journalism, Media and Culture.

https://datajusticelab.org/

Data & Society:

Advances public understanding of the social implications of data-centric technologies and automation.

https://datasociety.net/

DigiLabour:

The aim is to produce and share research about the world of work and digital technologies. The focus is on themes of platform work, platform cooperativism, self-management and alternative work arrangements; collective organisation of workers and digital technologies; algorithms and artificial intelligence at work; datafication and surveillance at work.

https://digilabour.com.br/

EthicsNet:

A community with the purpose of experimenting with different potential techniques to create datasets – examples of nice behaviours (such as social norms), to help socialise A.I.

https://www.ethicsnet.org/

FairWork Project:

Based at the Oxford Internet Institute and the WZB Berlin Social Science Center. The project evaluates the working conditions of digital platforms and ranks them based on five principles of fair work. It currently operates in 30 countries across 5 continents, working with workers, platforms, advocates and policymakers to envision and build a fairer future of work.

https://fair.work/en/fw/homepage/

FemLab Project:

A researcher activist cooperative that seeks to envision and enact how digital platforms may be optimised to enhance self-actualisation, representation, and collectivisation in a changing and increasingly precarious market and society. The project takes a worker-centred and feminist approach in the design and deployment of new media tools to align the future of work with human dignity, creativity, and pleasure.

https://femlab.co/

Hiperderechos

A non-profit civil society association dedicated to investigating, facilitating public understanding and promoting respect for rights and freedoms in digital environments.

https://hiperderecho.org/

Lavits (Rede Latino Americana de Estudos Sobre Vigilância, Tecnologia e Sociedade)

Aims to foster exchanges between Latin American researchers, activists and artists who focus on the connections between surveillance, technology and society.

https://lavits.org/a-lavit-pt/

Nesta:

The UK's innovation agency for social good. Nesta designs, tests and scales new solutions to society's biggest problems with the aim of changing lives for the better.

https://www.nesta.org.uk/

UNICRI Centre for Artificial Intelligence and Robotics:

Located in the Hague, UNICRI focuses on Goal 16 of the 2030 Agenda for Sustainable Development Agenda, that is centred on promoting peaceful, just and inclusive societies, free from crime and violence.

https://unicri.it/

United Tech and Allied Workers (UTAW):

A branch of the Communication Workers Union which seeks to represent and fight for workers interests in the tech industry.

https://utaw.tech/about/

The AI Now Institute:

The AI Now Institute aims to produce interdisciplinary research and public engagement to help ensure that AI systems are accountable to the communities and contexts in which they are applied.

https://ainowinstitute.org/

Witness.org:

Witness is a leader of a global movement that uses video to create human rights change. The organisation has led innovative initiatives on deep fakes, media manipulation and generative AI.

https://www.witness.org/

References

1 Check the Project A-Z of AI. The A-Z guide is a collaboration between the Oxford Internet Institute (OII) and Google - **https://atozofai.withgoogle.com/intl/en-US/artificial-intelligence/**

2 Verdegem, P. (2021, Ed.) AI for Everyone? Critical Perspectives. London: University of Westminster Press. **https://www.uwestminsterpress.co.uk/site/books/e/10.16997/book55/**

3 See the work of these scholars:
   Aguilar, Y.E. (2020). A Modest Proposal to Save the World. *The Rest of World.* Retrieved from: **https://restofworld.org/2020/saving-the-world-through-tequiology**
   Ali, S.M. (2019). "White Crisis" and/as "Existential Risk," or the Entangled Apocalyptism of Artificial Intelligence, *Zygon: Journal of Religion & Science*, 54(1), 207-224.
   Aouragh, M. and Chakravartyy, P. (2016). Infrastructures of Empire: Towards a Critical Geopolitics of Media and Information Studies. *Media, Culture & Society*, 38(4), 559-575.
   Brevini, B (2022) Is AI Good for the Planet? Cambridge: Polity.
   Ricaurte, P. (2022). Ethics for the Majority World: AI and the Question of Violence at Scale. *Media, Culture & Society*, 44(4), 726-745.

4 See Google.org Impact Challenge: Tech for Social Good - **https://impactchallenge.withgoogle.com/techforsocialgood/about**

5 See - Berendt, M. (2019). AI for the Common Good?! Pitfalls, Challenges, and Ethics Pen-Testing. Paladym, Journal of Behavioral Robotics, 10, 44-65.

6 Please refer to this definition - Floridi, L., Cowls, J., King, T.C. *et al*. How to Design AI for Social Good: Seven Essential Factors. *Sci Eng Ethics* 26, 1771–1796 (2020).
   https://doi.org/10.1007/s11948-020-00213-5

7 See: Freire, P. (1972). Pedagogy of the Oppressed. New York: Continuum / Freire, P and Macedo, D (1995) A Dialogue: Culture, Language, and Race. Harvard Educational Review 65(3): 379-390./
   Fals Borda, O (1987) The application of participatory action-research in Latin America. International Sociology, 2(4): 329-347 / Fals Borda, O (2003) Ante la crisis del país: Ideas acción para el cambio
   (1st. Ed.), Bogotá, Panamericana.

8 See - Alencar, A and Camargo, JF (2022) Stories of Migration: Exploring the Links Between Emotions and Technologies in the Narratives of Venezuelan Refugees in Brazil.
   doi: 10.1007/978-3-030-94122-2_6

Principal Investigator:

Dr Andrea Medrado

Senior Lecturer, School of Media and Communication

Communication and Media Research Institute (CAMRI)

University of Westminster

a.medrado@westminster.ac.uk

Co-Investigator:

Dr Pieter Verdegem

Reader, School of Media and Communication

Communication and Media Research Institute (CAMRI)

University of Westminster

p.verdegem@westminster.ac.uk