Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence (IJCAI-22)
Special Track on AI for Good (Projects)

A Norm Optimisation Approach to SDGs: Tackling Poverty by Acting on Discrimination

Georgina Curto[2,3,4,*], **Nieves Montes**[1,*], **Flavio Comim**[3,5], **Nardine Osman**[1] and **Carles Sierra**[1]
[*]Both authors have equally contributed to this paper
[1]Artificial Intelligence Research Institute, IIIA-CSIC, Barcelona, Spain
[2]University of Notre Dame, Notre Dame, USA
[3]IQS School of Management, Ramon Llull University, Barcelona, Spain
[4]EINA, Autonomous University of Barcelona, Spain
[5]Land Economy and VHI, University of Cambridge, England
{nmontes,nardine,sierra}@iiia.csic.es, {georginacurtor,flavio.comim}@iqs.url.edu

Abstract

Policies that seek to mitigate poverty by acting on equal opportunity have been found to aggravate discrimination against the poor (aporophobia), since individuals are made responsible for not progressing in the social hierarchy. Only a minority of the poor benefit from meritocracy in this era of growing inequality, generating resentment among those who seek to escape their needy situations by trying to climb up the ladder. Through the formulation and development of an agent-based social simulation, this study aims to analyse the role of norms implementing equal opportunity and social solidarity principles as enhancers or mitigators of aporophobia, as well as the threshold of aporophobia that would facilitate the success of poverty-reduction policies. The ultimate goal of the social simulation is to extract insights that could help inform and guide a new generation of policy making for poverty reduction by acting on the discrimination against the poor, in line with the UN "Leave No One Behind" principle. An "aporophobia-meter" will be developed and guidelines will be drafted based on both the simulation results and a review of poverty reduction policies at regional levels.

1 Problem Statement

Traditional poverty reduction policies have proved ineffective in the last decades. Despite the enormous efforts devoted to the redistribution of wealth, 700 million people, or 10% of the global population, still live in extreme poverty and evidence suggests that global poverty could increase by as much as 8% as a result of COVID-19 crisis, according to the United Nations (UN). The principles of equal opportunity and solidarity are key pillars of our welfare states and have been the main political answer to reduce poverty in terms of distributive justice. However, the rhetoric of equal opportunity has also been associated with the stigmatisation of the poor and the uneducated, who are considered blameful for not climbing up the social ladder [Sandel, 2020].

In an era of increasing inequality, the well-known "elephant curve of global inequality 1980 - 2020" [Chancel and Piketty, 2021] illustrates that the top 1% income distribution captured 23% of total world growth vs 9% for the bottom 50%. Meanwhile, the poor are told that they have the opportunity to prosper if they study at a good university and work hard [Mounk, 2017]. However, meritocracy has not worked as expected. Only 7% of United States citizens from the 20% lower rents reach the 20% top rents within their lifetimes [Chetty *et al.*, 2014]. Results are not better in many European countries (Germany presents even lower social mobility than the US [OECD, 2018]). The difference, however, is that Europeans tend to underestimate social mobility whereas in the US social mobility is overestimated [Alesina *et al.*, 2018].

It needs to be clarified that, even if policy makers tried their best to create an atmosphere close to perfect equal opportunity, there is no such thing since, from the moment of birth, individuals are exposed to different environments [Fishkin, 2014]. Furthermore, the rhetoric of equal opportunity constitutes an obstacle to pass and implement policies aimed at reducing poverty for the so-called "undeserving poor" [Everatt, 2008], forcing policy makers to determine which poor are victim of the circumstances ("luck egalitarism"), and therefore deserving aid, and which are responsible for their poverty [Anderson, 1999]. As an undesired effect, a social stigma associated with the poor, aporophobia [Cortina, 2017], is aggravated by blame. This has important psychological consequences for the well-being of people in need, who are avoided, discriminated or even attacked as a result of this socially-shared prejudice which leads to resentment, insecurity and self-hate.

The study of discrimination against the poor and its effect on poverty reduction has not received the attention it deserves in the literature. The Spanish philosopher Adela Cortina coined the term *aporophobia* to describe the rejection of the poor [Cortina, 2017], and the concept was included in the Spanish legal framework as an aggravating factor for hate crimes in 2021 [Boletín Oficial del Estado, 2021]. As an example of the ubiquity of discrimination against the poor that concerns the AI community, [Curto *et al.*, 2022] conducted a pioneer study informing about bias against the poor

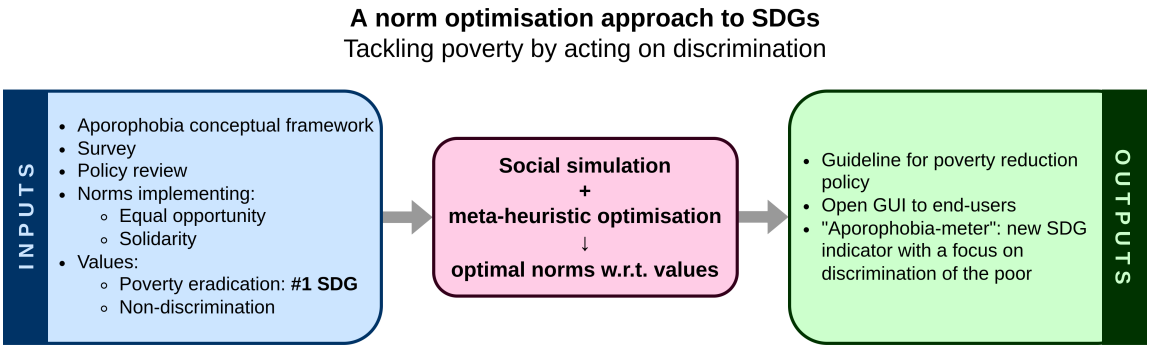


Figure 1: Collaboration proposal outline.

in Word2Vec and GloVe embeddings by using word vector representations. However, studies are still required to provide evidence on whether aporophobia hinders the success of policy reduction policies and can be considered an obstacle for the achievement of the first UN Sustainable Development Goal (SDG), poverty eradication.

Since testing the effectiveness of changes in policy and the resulting outcomes in real-life scenarios is an unfeasible process with great social and ethical repercussions, social simulation models emerge as powerful tools that, if well-formulated, can help assess the impact that new regulations have on a community. A recent example on the use of agent-based models to inform policy making concerns the COVID-19 pandemic and the simulation of non-pharmaceutical interventions on the evolution of infections.

In this context, this multidisciplinary project aims to answer the following question: (1) what is the impact of *prescriptive norms* (aka policy measures) related to equal opportunity and fiscal solidarity on aporophobia at a macro level?; and (2) to what extent does aporophobia at a micro level influence the effectiveness of poverty reduction policies?

We will study this problem in the context of *parametric norms*, where solutions for optimal levels of equal opportunity and redistribution of resources can possibly be obtained using search and optimisation techniques, such as meta-heuristics. Previous work on agent-based modelling for the evaluation and optimisation of normative systems of income transfer (i.e. tax collecting and redistribution) exists in the literature [Sallila, 2010]. However, it tackles poverty reduction from a merely resource redistribution perspective. As detailed in our introduction, the resentment created by such measures often leads to discrimination against the very people whom these policies are supposed to help, unfortunately rendering them ineffective.

The main innovation of our project, then, consists in the introduction of *aporophobia* both as a macro indicator that the norms in place ought to minimise, and as an individual attitude towards the acceptance of such norms. Prior to that, we will develop a conceptual framework and gather empirical data on aporophobia, based on the state of the art studies on prejudice, discrimination and bias [Pettigrew, 2021; Paolini *et al.*, 2021]. We expect that this work will be a key addition to traditional poverty reduction models.

2 Societal Benefits and Target SDGs

Assessing the impact of aporophobia on poverty levels opens a completely new path to tackle poverty reduction, from the point of view of non-discrimination measures and public awareness. Poverty could therefore be mitigated not only by aiding the poor through redistributive justice, but also by mitigating the existing discrimination.

Provided we succeed in modelling the relationship between equal opportunity, aporophobia and poverty, social simulation models will allow to investigate the role of discrimination in poverty reduction. We will define guidelines and indicators to mitigate poverty that will possibly include tackling aporophobia, in the case that aporophobia is found to be a main barrier towards poverty reduction. The resulting societal benefits can be grouped into the following categories:

- **Poverty reduction – number 1 SDG**: We expect to provide evidence that aporophobia is a brake for poverty reduction policies. When the poor are considered responsible for their fate, poverty reduction policies may be not be well received by the public, compromising their effectiveness. The work developed in this project will seek to quantify the effectiveness of poverty reduction policies as a function of the aporophobic levels of the population and provide recommendations for policy-making taking these results into account.
- **Reduce resentment**: An important characteristic of poverty goes beyond the lack of material goods. As Nobel Laureate Amartya Sen states, a person is not poor when he or she can carry out a meaningful life with dignity [Sen, 2004]. However, the rhetoric of equal opportunity in a time of growing inequality can lead to the frustration of some of the population in need [Hochschild, 2018]. As a result of the guidelines and indicators provided as outputs of this study, awareness campaigns can be put in place to mitigate the social stigma associated to the poor. This constitutes a new paradigm on poverty reduction approaches, since the focus is not only on the poor, but also on the non-poor and society in general.
- **Enhanced sense of community**: This study will provide the optimal equilibrium between norms implementing equal opportunity and social solidarity. Having equal

opportunity as the main answer to poverty creates a high level of competition among citizens, [Mounk, 2017; Sandel, 2020], enhanced by increasing inequality in the era of globalisation [Osnos, 2014]. When other non-competitive solutions are in place in terms of contributive justice, such as recognising the value of jobs independently of the salaries, individuals perceive social recognition and part of a community [Smith and Deranty, 2011].

3 Goals and Methods

The goals that this collaboration aims to achieve involve two distinct but complementary disciplines: the social sciences (in particular philosophy, psychology and welfare economics), and computer science and AI.

From the psychological and philosophical perspectives, it is of utmost importance to, early on, establish a clear working conceptual framework for aporophobia and its relationship to other relevant concepts such as social values, norms and attitudes. To do so, a literature review of the state of the art research on the specific topics of prejudice, discrimination, stigma and bias will be conducted. Additionally, in order to counteract the lack of specific psychological research on the topic, the authors intend to design, conduct and analyse their own aporophobia survey.

From a computational perspective, the main objective of this collaboration is the formulation and implementation of an agent-based simulation to study the interdependence between social values, prescriptive norms and aporophobia. Such a program could help us quantify the relationship between these three components, in particular whether aporophobia acts as an inhibitor of effective poverty reduction policy making.

Methodologically, we intend to follow the framework originally presented by the authors in [Montes and Sierra, 2021; Montes and Sierra, 2022] to develop such a social model. In general terms, the method consists of the following steps. First, the state features that are relevant in the domain where the simulation is focused are to be defined. Provided that we are interested in poverty from a multidimensional perspective, potential features will look not only for resource deprivation, but also health and education failures, among others. These definitions will require a close collaboration between the AI and the the social sciences team members.

Second, the prescriptive norms (aka the regulations) that govern the society of agents have to be defined as well as the optimisable parameters they are tied to, i.e. the *normative parameters*. The union of all the normative parameters, together with their bounds and constraints, provides the space where a search algorithm will look for the optimal levels. In this project, we intend to focus on norms implementing *social solidarity* and *equal opportunity*. Therefore, it will be necessary to formulate how these norms operate and what is their impact on the behaviour of the agents.

Third, the semantics of the *values* of interest have to be defined. We understand values as abstract general concepts, whose meaning is grounded in a particular context (a socio-economic simulation in this case) by a particular goal or function. Values, hence, are operationalised through these goals and functions, which serve to evaluate the states of the system and the achieved outcomes. The semantics of such values will provide us with the optimisation target as the *alignment* of a candidate normative system with respect to the value of interest. For this project, we are interested in values related to the erradication of poverty and the reduction of aporophobia. For example, a wide variety of indicators for economical equality are available [Cowell, 2011], of which the authors will need to discern and pick the most suitable ones or possibly come up with their own indicators with a stronger focus on discrimination.

Fourth and final, a suitable AI technique has to be chosen, considering the scale and computational requirements of the resulting social model, to automatically perform the search over the space of normative parameters with the target on aporophobia and poverty reduction. We will be looking for the optimal levels of solidarity and equal opportunity that would maximise the promotion of the values of focus, *i.e.* "minimum poverty" and "maximum fairness" (understood as an ideal desired value of non-discrimination against the poor). In order to have a meaningful assessment of the effect of a candidate normative system, a sampling technique over several runs of the model, such as Monte Carlo sampling, is necessary, regardless of the search algorithm of choice. In the past, the authors have settled for a Genetic Algorithm (GA) to conduct the search [Montes and Sierra, 2021]. Although we do not discard, right off the bat, other optimisation techniques, in particular meta-heuristics ones such as Simulated Annealing, many of the reasons that made us settle for GAs in our previous work are applicable in this project, namely their versatility for optimisation over continuous, discrete and hybrid (both continuous and discrete variables) domains [Luke, 2013]. It is worth noting that GAs are also very suited for parallelisation, as the recombination of parent normative systems for the creation of the new generation can be easily distributed over several computing nodes. This is a clear point in favour of sticking to this class of search algorithms, as we foresee that the scale of a meaningful model, that is able to reflect with some degree of fidelity real-life scenarios, will require intensive computational resources.

Additionally, a major extension to the norm optimisation framework in [Montes and Sierra, 2021] needs to be developed in order to account for the aporophobic attitudes of the agents in the system. As a preliminary idea, we contemplate modelling aporophobia at the micro level as an internal psychological agent construct linked to the ranges of the normative parameters an individual is willing to accept. This, in turn, conditions their willingness to abide by the established norms. We would like to emphasise, however, that the computational construct to introduce aporophobia into the model will be reviewed and refined after the authors have finalised the conceptual framework of aporophobia and obtained empirical evidence of the nature of the phenomenon through surveys. We expect that this prior revision will make the agent-based model more sound and better grounded in the state of the art of social psychology.

The final objective that the collaboration pursues is the closure of the feedback from the computational results to the so-

cial sciences. The ultimate goal of the social simulation is to extract insights that could help inform and guide a new generation of policies for poverty reduction. To achieve this goal, a policy review of regional legislation (including a comparative analysis of the Global North and the Global South) is to be conducted. Comments, suggestions and guidelines will be drafted based on the simulation results. If, hypothetically, the simulation results conclude that no level of aporophobia should be allowed in order for poverty-reduction measures to be effective, recommendations would be made on targeting the aporophobia levels of the population through communication campaigns. Other insights from the simulation results would need to be translated back into the realm of the social sciences in terms of other measures, possibly targeting the adequate balance between social solidarity and equal opportunity.

As an final touch to the project, and with the intention to reach as many policy makers and social science scholars as possible, we will develop a graphical user interface (GUI) to the social simulation or a lightweight version of it. With a GUI, anyone could interactively switch parameter values or define what values are they interested in promoting and examine the effect that those changes would have on a community of agents. Such a step would make the technical aspects of this research much more accessible to the people that should extract insights and act upon it.

4 Challenges and Risks

We identify five main risks to be tackled during the execution of the proposal, which we itemise alongside with their mitigating measures in Table 1. These can be categorised into two classes. The first concerns the quality of the agent-based model (rows one, two and three in Table 1). In order to ensure that the formulated model is sound, unbiased and produces relevant insights, it will be grounded on the state-of-the-art on discrimination, poverty, equal opportunity and social solidarity literature, as well as the incipient literature on aporophobia [Curto *et al.*, 2022]. In the context of this collaborative project, the authors will also develop their own survey to obtain empirical data on the levels of aporophobia in human subjects. However, basing the agent-based model on data extracted from a small subset of the world population (namely people from western countries or the Global North) would render the model biased towards the necessities of these societies and the policy recommendation derived from the results irrelevant to a wide range of audiences. To avoid this, the model needs to be adaptable to local characteristics which, in turn, will require conducting the survey on subjects from many different backgrounds and developing the simulation in a modular way.

Another challenging task related to the formulation of the model is the definition of the semantics of values *minimum poverty* and *maximum fairness (non-aporophobia)*, which we want to embed into the system. This is a very important point since it defines the objectives that we would like to achieve in the simulations and will direct the recommendations derived from that. Relative and absolute concepts of poverty will be used, in line with the state-of-the-art literature on human development, considering also the impact of inequality. The analysis of poverty will be multidimensional and based on a lack of basic capabilities, as described by Nobel Laureate Amartya Sen, to avoid generating recommendations only based on reaching above poverty threshold incomes, which could have perverse effects.

The second category of risks to be tackled concerns the availability of resources to carry out the project, which are being mitigated as described in Table 1.

5 Evaluation Criteria

Once the agent-based model is implemented and the methodology is applied, a set of optimal norms will be obtained. These norms will implement a poverty reduction policy strategy and their effects will be quantified. In addition to these results, the authors have also developed a set of evaluation analytical tools to examine the resulting normative systems, optimised for poverty reduction, in depth.

The first of these evaluation tools consists in conducting a Shapley value analysis of the resulting optimal norms. This approach is grounded on the very well established field of co-operative game theory and considers that every individual norm (aka social solidarity and equal opportunity) is a member of a coalition, i.e. the normative system at large. Then, it is possible to apply the definition of Shapley value, taking as the worth of coalitions the *alignment* that these normative systems are able to achieve with respect to poverty reduction targets. Such a computation yields a quantitative evaluation on the importance of every individual norm when it comes to poverty reduction. This is a very informative metric that helps discern the mechanisms by which the optimised normative systems are achieving their targets.

The second of these evaluation tools regards the compatibility of poverty reduction with other values that policymakers and scholars might deem relevant, such as social mobility. In our previous research, we have encountered that normative systems that are highly optimised for some values can be very oblivious for others [Montes and Sierra, 2022]. Additionally, if the users of the model deem it interesting, it is also possible to perform optimisation searches with the target being not a particular value such as poverty reduction, but the *compatibility degree* among several values. This type of search will produce the normative system that is the best compromise for a set of different values.

However, this computational evaluation is not the whole story, as we will need to check if humans accept the norms that have been obtained, no matter what are their initial levels of aporophobia and whether, after analysing them, their levels of aporophobia get reduced. In order to perform this evaluation, surveys will be constructed to measure and follow up on the levels of aporophobia, before and after exposing the subjects to the reading and analysis of the norms. A well-known problem of questionnaires that seek to measure ethical topics in the population is bias, since individuals do not tend to be completely honest. This will been taken into account in the questionnaire design according to the state-of-the-art on survey design on moral topics. [Greenwald *et al.*, 2009]. It is imaginable that if the results leave room for improvement,

Risk description	Mitigating actions
Ungrounded social model	The model will be based on the state-of-the-art on discrimination and poverty, as well as empirical data on aporophobia obtained from the the author's own survey. In addition, the model will be based on a methodology that has been proven successful in the past [Montes and Sierra, 2021; Montes and Sierra, 2022].
Biased social model	The survey to obtain data will be used at a global level and, when used regionally, the resulting recommendations will be tailor-made by adapting the model to local characteristics. Demographic data will be collected through the survey to ensure the representativity of income levels, education and professional backgrounds. Especial attention will be paid to the representation of historically discriminated groups.
Irrelevant results	A review of the poverty reduction policy framework at a regional and global level (including the Global North and the Global South) will be performed so that the guidelines resulting from the project provide added value to international and regional NGOs and government officials.
Lack of financial resources	The two direct costs of the project are human resources and computational equipment. These are available to the existing team since this project has been incorporated as an internal objective. Additionally, the team has applied to obtain funding through the call of Fundació La Caixa for Innovative Projects in the field of social sciences.
Lack of computating power	The IIIA team has the support of an HPC service and access to the Ars Magna cluster. In order to leverage these computational resources, a search strategy amenable to parallelisation, such as the Genetic Algorithms discussed in Section 3, will be implemented.

Table 1: Potential risks that will encountered during the execution of the process and the corresponding mitigating actions.

further iterations of modelling, optimisation and evaluation would be run.

6 Collaboration Mechanisms, Timeline and Outputs

This proposal is markedly multidisciplinary and has a global perspective. The know-how of a diversity of disciplines and AI techniques is leveraged to analyse and propose a new approach to poverty reduction, feasible through computational simulation and resulting into recommendations for poverty reduction policies both at a regional and international levels. Nevertheless, in order to achieve this goal, it is essential to define an accurate conceptual framework and obtain the appropriate inputs to feed the simulation. The international team presenting the project idea is composed of social science specialists in the topic (welfare economics and aporophobia) as well as AI researchers (multi-agent systems, norms, metaheuristic optimisations) and welcomes inputs and participation from scholars in the areas of psychology, development economics, law, multiagent systems and optimisation techniques, who can surely provide valuable inputs.

The project will last for one calendar year and can divided into three main tasks: (1) formulation of a conceptual framework for aporophobia; (2) development of the social simulation model; and (3) obtaining results and extracting insights from it. Naturally, the tasks will be lead by the team members with more expertise, with the participation of researchers from a different background, (e.g. the AI team members will be in charge of leading the development of the social simulation model). In order to coordinate all the participants and achieve a cohesive understanding of the concepts that this project revolves around (namely, *aporophobia*, *prescriptive norms* and *values*), we intend to hold regular online meetings and cross-readings, where concepts should be clarified, progress reports will be presented and next steps forward will be discussed.

Besides these routine meetings and ongoing team work, we foresee three specially important meetings. The first one is the kick-off meeting, which has the objective to assign the first tasks, deliverables and define calendar deadlines. A second important meeting must also happen once the social science team members can provide a first draft of the conceptual framework on aporophobia. Both teams will discuss the proposal in order to define how to introduce it as modular components of the social simulation. The third important meeting will happen once the results from the social simulation are available. Both teams will comment on the results in order to extract insights, the potential need for iterations (that might imply modifications in the conceptual model) and draft policy-making recommendations.

Regarding scientific production, we expect the following outputs from the project:

1. A first scholar paper will detail the conceptual framework of the project (i.e. how the economic, psychological and ethical aspects of aporophobia relate to the norms and values to be used within the social simulation). This paper will also include empirical results on aporophobia levels measured on individuals obtained from the foreseen survey. The raw data from the results of such survey will be made publicly available, conveniently anonymised, through a dataset platform.
2. A conference paper, preferably for IJCAI-ECAI 2023 or 2024, containing the framework for the synthesis of normative systems, with an extension to incorporate aporophobia both at a macro and a micro levels, the results and the analysis of the simulation output. Of course, all the code developed during this stage will be properly

documented and made available in open source for other researchers and anyone in the general public to examine.

3. A policy guidelines document, informed by the results of the social simulation, to guide on new paths for poverty reduction policies based on the mitigation of aporophobia. Although we do not rule out the possibility of journal submission, we intend to keep all versions of this document open and accessible to anyone.

4. If possible, according to time and financial constraints, a graphical user interface with an interactive version of our social simulation model. A public repository with installation and running instructions will be made available on platforms such as GitHub and AI4EU.

7 Long-term Impact on the SDGs

This project aims at opening a completely new path to tackle poverty reduction at a macro-global, meso-national and micro-individual levels, which is UN's #1 SDG (poverty eradication) and is clearly related to the other 16 SDGs, due to the multidimensional nature of poverty, and in particular to #2 SDG (zero hunger), #8 SDG (decent work and economic growth) and #10 SDG (reduced inequalities). The proposed work is disruptive by informing about the optimal levels of norms related to "equal opportunity" and "social solidarity" to attain the values of "minimum poverty" and "maximum fairness" (non-aporophobia), while at the same time considering discrimination factors. The project is also groundbreaking since it allows to explore alternative approaches to poverty reduction within an original AI simulation context, based on existing data and providing recommendations that aim to be potentially applied on a real-life scenarios, both regionally and globally.

It has been suggested that the discrimination against the poor (aporophobia) could have an impact on poverty at different levels. From a macro-international perspective, the developing countries are considered responsible for their fate, instead of working towards a global equilibrium in areas such as international commerce, cooperation among countries and financial markets. At a meso-national level, aporophobia hinders the effective implementation of poverty reduction measures. Finally, at a micro-personal level, the self-depreciation of the poor is an additional obstacle to improve their economic situation.

By providing evidence that aporophobia constitutes an obstacle for poverty reduction, the study opens the opportunity to a completely new set of measures to reduce poverty, acting on the socially shared prejudices towards the poor. The focus of the problem (and the solution) would be not only on the poor, but also on the non-poor and the society as a whole.

The "aporophobia-meter" will allow to measure and follow up the evolution of aporophobia, enabling policy makers to relate it with poverty levels throughout time. A long-term goal of the study is to encourage a virtuous circle where less discrimination against the poor lead to a higher effectiveness of poverty reduction policies at global and regional levels.

Ethical Statement

The project will comply with the Trustworthy AI principles described in the Ethics Guidelines for Trustworthy AI [HLEGAI, 2019], the White Paper on AI [European Commission, 2020] and the proposed EU Artificial Intelligence Act [European Union, 2021].

The project proposal has been submitted to the Research Ethics Committee (REC) of the Universitat Autònoma de Barcelona (UAB) and the Ramon Llull University, following an Ethics Assessment. The main ethics question is in relation to Personal Data. This project will adhere to the EU General Data Protection Regulation (GDPR) and the Spanish Organic Law 3/2018 of 5 December 2018 on Personal Data Protection and guarantee of digital rights. While the data that can be gathered from surveys can contain demographic information, the results and conclusions will be redacted in an anonymous manner. Every step will be taken to ensure that the results do not negatively affect any particular person or entity.

References

[Alesina *et al.*, 2018] Alberto Alesina, Stefanie Stantcheva, and Edoardo Teso. Intergenerational mobility and preferences for redistribution. *American Economic Review*, 108(2):521–554, February 2018.

[Anderson, 1999] Elizabeth S. Anderson. What is the point of equality? *Ethics*, 109(2):287–337, jan 1999.

[Boletín Oficial del Estado, 2021] Boletín Oficial del Estado. Circular 7/2019, de 14 de mayo, de la fiscalía general del estado, sobre pautas para interpretar los delitos de odio tipificados en el artículo 510 del código penal. BOE-A-2019-7771, 2021. Ministerio de la Presidencia, Relaciones con la Cortes y Memoria Democrática (Gobierno de España).

[Chancel and Piketty, 2021] Lucas Chancel and Thomas Piketty. Global income inequality, 1820-2020: The persistence and mutation of extreme inequality. *Journal of the European Economic Association*, 2021.

[Chetty *et al.*, 2014] Raj Chetty, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. Where is the land of opportunity? the geography of intergenerational mobility in the united states. *The Quarterly Journal of Economics*, 129(4):1553–1623, sep 2014.

[Cortina, 2017] Adela Cortina. *Aporofobia, el rechazo al pobre : un desafío para la democracia*. Estado y Sociedad. Ediciones Paidós, 2017.

[Cowell, 2011] Frank Cowell. *Measuring Inequality*. Oxford University Press, March 2011.

[Curto *et al.*, 2022] Georgina Curto, Mario Jojoa Acosta, Flavio Comim, and Begoña Garcia-Zapirain. Are AI systems biased against the poor? A machine learning analysis using Word2vec and GloVe embeddings. Forthcoming, 2022.

[European Commission, 2020] European Commission. White paper on artificial intelligence: a European approach to excellence and trust. White Paper COM(2020) 65 final, European Commission, Brussels, February 2020.

[European Union, 2021] European Union. Proposal for a regulation of the European Parliament and of the Council – laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts, 2021. COM/2021/206 final.

[Everatt, 2008] David Everatt. The undeserving poor: Poverty and the politics of service delivery in the poorest nodes of south africa. *Politikon*, 35(3):293–319, dec 2008.

[Fishkin, 2014] Joseph Fishkin. *Bottlenecks*. Oxford University Press, feb 2014.

[Greenwald *et al.*, 2009] Anthony G. Greenwald, T. Andrew Poehlman, Eric Luis Uhlmann, and Mahzarin R. Banaji. Understanding and Using the Implicit Association Test: III. Meta-Analysis of Predictive Validity. *Journal of Personality and Social Psychology*, 97(1):17–41, jul 2009.

[HLEGAI, 2019] HLEGAI. *Ethics guidelines for trustworthy AI*. Publications Office, 2019. High-Level Expert Group on AI – European Commission and Directorate-General for Communications Networks, Content and Technology.

[Hochschild, 2018] Arlie Russell Hochschild. *Strangers In Their Own Land*. Ingram Publisher Services, March 2018.

[Luke, 2013] Sean Luke. *Essentials of Metaheuristics*. Lulu, second edition, 2013. Available for free at http://cs.gmu.edu/∼sean/book/metaheuristics/.

[Montes and Sierra, 2021] Nieves Montes and Carles Sierra. Value-guided synthesis of parametric normative systems. In *Proceedings of the 20th International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '21, page 907–915, Richland, SC, 2021. International Foundation for Autonomous Agents and Multiagent Systems. (Best paper award finalist).

[Montes and Sierra, 2022] Nieves Montes and Carles Sierra. Synthesis and properties of optimally value-aligned normative systems. Under review, 2022.

[Mounk, 2017] Yascha Mounk. *The Age of Responsibility*. Harvard University Press, May 2017.

[OECD, 2018] OECD. *A Broken Social Elevator? How to Promote Social Mobility*. OECD, 2018.

[Osnos, 2014] Evan Osnos. *Age of Ambition: Chasing Fortune, Truth, and Faith in the New China*. Farrar, Strauss and Giroux, May 2014.

[Paolini *et al.*, 2021] Stefania Paolini, Fiona A. White, Linda R. Tropp, Rhiannon N. Turner, Elizabeth Page-Gould, Fiona K. Barlow, and Ángel Gómez. Intergroup contact research in the 21st century: Lessons learned and forward progress if we remain open. *Journal of Social Issues*, 77(1):11–37, mar 2021.

[Pettigrew, 2021] Thomas F Pettigrew. Advancing intergroup contact theory: Comments on the issue's articles. *Journal of Social Issues*, 77(1):258–273, mar 2021.

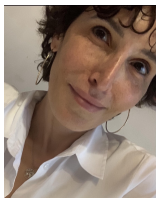
[Sallila, 2010] Seppo Sallila. Using microsimulation to optimize an income transfer system towards poverty reduction. *Journal of Artificial Societies and Social Simulation*, 13(1), 2010.

[Sandel, 2020] Michael J Sandel. The tyranny of merit. *Penguin Random House*, 2020.

[Sen, 2004] Amartya Sen. *Development as Freedom*. Oxford University Press, January 2004.

[Smith and Deranty, 2011] Nicholas Smith and Jean-Philippe Dr. Deranty, editors. *New Philosophies of Labour: Work and the Social Bond*. Brill, November 2011.

About the Authors



Georgina Curto has conducted this work as a PhD student at the joint program of universities Ramon Llull (IQS School of Management), Comillas (ICADE) and Deusto, under the supervision of Prof Flavio Comim. Her research has developed in the area of AI Ethics and focuses on bias and discrimination against the poor. Before the PhD, she obtained her Research Masters with a final thesis on the topic of Design Thinking at Ramon Llull University and a Master in Business Administration at Instituto de Empresa (IE). She currently teaches Strategic Design at Universitat Autònoma de Barcelona (EINA) and she will be joining the Technology Ethics Center at the University of Notre Dame as a postdoctoral fellow.

Nieves Montes is a PhD student at the Artificial Intelligence Research Institute (IIIA-CSIC) under the supervision of Prof Carles Sierra and Dr Nardine Osman. Her research involves the study and development of autonomous agents who are able to act in social context, particularly by leveraging norms, values and awareness of others. Prior to embarking on her PhD, she obtained her Masters at the Autonomous University of Barcelona with a thesis on value alignment in multiagent systems. Her work presented presented at the International Conference of Autonomous Agents and Multiagent Systems (AAMAS) 2021 was selected as a finalist for the best paper award.

Flavio Comim is the Head of Chair of Ethics at Universitat Ramon Llull University (IQS). He obtained his PhD in Economics at the University of Cambridge (1998), his Mphil in Economics at the University of Cambridge (1994) and his MSc at the University of Sao Paulo (1993). His research lines are on the topics of education and human development, impact evaluation of corporate social investment, the capability approach, poverty, inequality and the common good.

Flavio held the position of Senior Economist for Brazil's United Nations Development Programme (UNDP) (2008-2010) and has been a consultant for UN agencies such as World Health Organisation (WHO), United Nations Environment Programme (UNEP), Food and Agriculture Organisation (FAO), United Nations Educational, Scientific and Cultural Organisation (UNESCO) and UNDP-NY.

Nardine Osman is a Senior Scientist at the Artificial Intelligence Research Institute (IIIA) of the Spanish National Research Council (CSIC). She obtained her PhD (2008) and Master (2003) in Informatics from the University of Edinburgh, UK. In recent years, Nardine's research has focused on ethical issues of AI, especially the issue of value engineering. She has led lately a number of projects on related topics (Spanish funded CIMBVAL, RecerCaixa funded AppPhil, and EU funded WeNet). Her background expertise is in agreement technologies, norms and normative systems, and formal methods in general. She is a member of the JAIR Editorial Board, and was the vice president of EURAMAS until recently.

Carles Sierra is Director of the Artificial Intelligence Research Institute (IIIA) of the Spanish National Research Council (CSIC) located in Barcelona. He is the President of EurAI. He has been contributing to Artificial Intelligence research since 1985 in the areas of Knowledge Representation, Auctions, Electronic Institutions, Autonomous Agents, Multiagent Systems and Agreement Technologies. He is or has been a member of several editorial boards of journals, including AIJ and JAIR, two of the most prestigious generalist journals, and was the editor in chief of the JAAMAS journal, specialised in autonomous agents. He organised IJCAI, the most important international artificial intelligence conference in 2011 in Barcelona and was the President of the IJCAI Program Committee in 2017 in Melbourne. He is a Fellow of the European Association of AI, EurAI, and recipient of the ACM/SIGAI Autonomous Agents Research Award 2019.