THE IMPLICATION OF AI
IN SOCIAL WELFARE SYSTEMS:
POTENTIAL RISKS AND PREVENTION MEASURES

By

Akmaljon Akhmedjonov

Submitted to
Central European University
Department of Public Policy

In partial fulfillment for the degree of
Master of Arts in International Public Affairs

Supervisor:

Cameran Ashraf

Vienna, Austria
2023

Author's Declaration Form

I, the undersigned, *Akmaljon Akhmedjonov*, candidate for *Master of Arts in International Public Affairs*, declare herewith that thepresent thesis is exclusively my own work, based on my research.

All sources have been properly credited in the text, notes, and the bibliography. I declare that no unidentified and illegitimate use was made of the work of others, and no part of the thesis infringes onany person's or institution's copyright. Furthermore, I declare that no part of this thesis has been generated using artificial intelligence (ChatGPT).

I also declare that no part of the thesis has been submitted in this form as coursework for credits or toanother institution of higher education for an academic degree.

Date: May 29, 2023

Name: Akmaljon Akhmedjonov

Signature: 

Acknowledgments

I want to take a moment to thank all my professors, friends, family, and myself for the love and trust during this journey.

Abstract:

The use of Artificial Intelligence (AI) powered applications has become an integral ingredient in boosting efficiency and cutting costs both in the private and public spheres. Government institutions are becoming increasingly reliant on AI systems to detect fraud in tax, and corruption and make decisions regarding social welfare distribution. However, in the absence of nationally and internationally recognized regulatory institutions and laws, some countries are using AI for surveillance and control of their citizens which is a direct infringement of human rights and fundamental democratic values. To prevent further disruptions and misuse the EU AI Act proposal was introduced in April of 2021 which serves as a stepping stone to regulate AI systems in the EU market that might potentially help to resolve issues on preventing bias, unfairness, transparency, and accountability in all sectors including in social welfare that current paper is focused on. The present research aims to answer the following questions: (1) What are the implications of AI in general and the EU AI Act in social welfare? (2) What are the potential loopholes in the EU AI Act proposal which can harm human rights? And lastly how to balance the participation of different voices in the implementation process of AI applications in social spheres? During the research it was revealed that current regulatory institutions are not responsive to the risks caused by AI and several governments are misusing AI for the sake of efficiency over human rights.

TABLE OF CONTENTS

Author's Declaration Form ..... *ii*

Acknowledgments ..... *iii*

Abstract: ..... *iv*

Introduction *1*

1.1 Research objectives: **2**

1.2 Research problem: **3**

1.3 Research questions: **3**

1.4 Methodology **3**

1.5 Literature Review **4**

1.5.1 Disinformation 9

1.5.2 Involvement of Different Stakeholders 11

Chapter I EU AI Act and its possible guidelines on Social Welfare *14*

1.6 The indirect impact of AI on Social Welfare Decision-Making **16**

Chapter II: Risk of AI in Social Welfare *20*

1.7 Preventive Measures **22**

Chapter III: Discussions & Conclusion *24*

1.8 The Use of AI Software **26**

1.9 Regulation vs Progress **28**

Conclusion *29*

Bibliography **31**

LIST OF ABBREVIATIONS

ADM	Automated Decision Making
AI	Artificial intelligence
AMS	the Public Employment Service Austria (Arbeitsmarktservice)
EU	European Union
GAN	Generative Adversarial Networks
GDPR	General Data Protection Regulation
HLEG	High-Level Expert Group
OCI	Online Compliance Initiative
OECD	Organization for Economic Co-Operation and Development
SME	Small and medium-sized Enterprises
WHO	World Health Organization

Introduction

To solve the AI problem, rely on policy, not technology – Kate Crawford (one of the leading scholars on politico-social implications of AI)

In recent years the concept and phrase of Artificial Intelligence (AI) has dominated the headlines. Despite being overhyped, it is noticeable that the use of AI in the decision-making processes of critical sectors such as health, architecture, politics, and the economy has been exponentially rising. AI also like other technology has the potential to solve the complex challenges of mankind. However, the difference perhaps from traditional technologies such as printing machines or cars is that it is seemingly on steroids given its rapid growth and self-improving ability. Rather sudden growth automation comes with a new wave of issues concerning respecting human rights and in particular its ethical implications on the ground. If earlier AI was on the backend of operating systems, lately, it is also present in the front end of providing services such as making decisions on distributing welfare.

The recent developments in the sector also became handy for governments to "increase" efficiency and cut human costs to use algorithmic or machine learning systems in spheres like identifying fraud in tax, distribution of welfare, policing, and others. One of the reasons for the prevalent use of AI comes with a perception of solving challenges faced by humans to tackle human bias, racism, and other institutional inequalities, including building a better social welfare system. However, the reality presents a different picture where AI is not solving the issues, but instead copying and repeating human-made mistakes. As for the main theme of this paper understanding the implications of AI technology in social welfare requires recognizing its pivotal role within the current context of increasing constraints on social welfare provision, leading to a growing divide between the rich and poor. AI applications are playing and will

play a crucial role in determining who qualifies for social welfare, why certain individuals are eligible, and how welfare is distributed.

One of the purposes of this paper is indeed to analyze the front-end use of AI in social welfare which has not been tackled from the regulatory lenses. The reason to conduct such research comes from the economic turbulences due to COVID and wars where the number of people who live in extreme poverty has grown from 8.3 % in 2019 to 9.3% in 2020 which is equal to more than 700 million people worldwide.[1] According to the International Labor Organization currently, only 47% of the world's population have at least one social protection benefit while the rest 53% (around 4 billion people) do not have any protection provided by the state.[2] In the absence of ethical regulatory frameworks in the newly emerging tech sector, the provision of social protection will only decrease.

Many governments are effectively using AI to seek solutions for better social welfare but the question on the table is whether it is effective or not and what are the challenges ahead. Thus, one needs to ask how we should perceive the introduction of automation in social welfare systems. Can we solve issues regarding inequality with the help of AI and if yes, are the costs of the solution greater than its benefits?

1.1 Research objectives:

The objective of this research is to assess international legal frameworks/policies such as the EU AI Act proposal, UNESCO AI ethics guidelines, and cases of the implication of AI in social welfare in the context of Europe. The research will highlight the challenges and outcomes of using AI in social welfare with an emphasis on human rights, ethics, and the involvement of multiple actors in the decision-making process. Moreover, the paper will try to shed light on

[1] "Poverty," Text/HTML, World Bank, accessed April 30, 2023, https://www.worldbank.org/en/topic/poverty/overview.
[2] "More than 4 Billion People Still Lack Any Social Protection, ILO Report Finds," News, September 1, 2021, http://www.ilo.org/global/about-the-ilo/newsroom/news/WCMS_817653/lang--en/index.htm.

the understanding of the dilemma between progress and threats caused by AI in policy decision-making processes.

1.2 Research problem:

The research problem ultimately comes from the failure of the states to respond to the emerging challenges caused by AI in a timely manner. The technologies that are in place are outpacing legislative frameworks leaving an open loophole for an uncontrolled process that can potentially harm humans. Besides the lack of human capital (experts) who are willing to work for public institutions has become burdensome because the private sector is more lucrative both in terms of innovation and opportunities which leaves the legal & policy-making institutions to have no oversight over the application of AI in wrong purposes such as discrimination, cyberattacks or worse, online terrorism.

Lastly, the problem of applying AI in social welfare schemes is still a novel practice due to the absence of internationally agreed frameworks/guidelines on the use of AI. For these reasons, it is essential to consider the different scenarios together with the best practices for using AI for the well-being of humanity.

1.3 Research questions:

Essentially, the paper will try to assess the implication of AI in social welfare decision-making processes using both theoretical and practical angles by mainly answering these questions:

1. What is the implication of AI in social welfare?
2. What are the potential loopholes in the EU AI Act proposal which can harm human rights?
3. How to use the whole-of-society approach in using AI in decision-making processes?

1.4 Methodology

The methodology is an important component of research that seeks to investigate and comprehend the influence of AI on social welfare. This research evaluates possible loopholes

in the EU AI Act proposal within the context of applying AI in social welfare systems. The research will rely primarily on secondary qualitative data.

To gather information for the qualitative component, existing online expert interviews, podcasts, comparative legal documents (e.g., UNESCO and OECD AI Ethics Guidelines and EU AI Act), observations, and case studies will be used. This method allows for more in-depth knowledge of the experiences and viewpoints of those affected by AI in social welfare. By analyzing these documents, media reports, and research this paper will attempt to answer the research questions on the feasibility of introducing AI regulatory frameworks and their effect on citizen rights and freedoms along with possible threats caused by AI.

A possible limitation of this method used in this paper could be the lack of quantitative sources except for smaller statistical data on the effect of AI in decision-making processes. Moreover, the introduction of new concepts such as a whole society approach and threat modeling will help to bring ideas for potential policy improvements.

1.5 Literature Review

Even though AI is not a novel technology, its shining years started in the recent decade and thus the extensive research. This literature review is divided into three parts. The first part will assess how ethics is understood within the domain of using AI in social welfare decision-making. The second part will focus on different case studies where the use of AI is already in place together with policies and proposals such as the EU AI Act. And the third part will examine the literature on the whole-of-society approach and threat modeling on using AI in policy-making procedures both by private and public sector institutions.

Before going in-depth into the topic, it is essential to set up a clear understanding of the main concepts used in this research such as AI and social welfare. However, it is important to keep in mind that there is not a single agreed definition of AI globally. For instance, the EU High-

Level Expert Group on AI (AI HLEG) defines AI systems as "software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal."[3] In terms of technicalities, HLEG also explains that AI systems have the capability to employ either symbolic rules or acquire a numeric model, and they can further adjust their behavior by assessing the impact of their prior actions on the environment.

On the other hand, according to the OECD's definition, "AI is a general-purpose technology that has the potential to improve welfare and well-being of people, to contribute to positive sustainable global economic activity, to increase innovation and productivity, and to help respond to key global challenges."[4] It is a technology that is used in several sectors from manufacturing, banking, transportation, and security to healthcare. When it comes to the definition of social welfare or welfare it is understood as a state or set of institutions that play an important role in protecting the socio-economic well-being of its citizens which might include but is not limited to pensions, access to healthcare, or job provision.[5] This is rather a simplistic definition of social welfare, but it goes much deeper. There is often a misconception of welfare refers to helping the poor or disadvantaged but the welfare state firstly is not only about welfare nor about assisting the poor. It is instead about the provision of social rights and insurance, regulations regarding the economy, and financial action which benefits the middle class rather than the poor as well as employers. Moreover, the welfare state is not about state

[3] Fair Trials, "Automating Injustice: The Use Of Artificial Intelligence & Automated Decision-Making Systems In Criminal Justice In Europe," 2021, https://www.fairtrials.org/app/uploads/2021/11/Automating_Injustice.pdf.
[4] "The OECD Artificial Intelligence (AI) Principles," accessed May 1, 2023, https://oecd.ai/en/ai-principles.
[5] "Welfare State | Britannica," April 18, 2023, https://www.britannica.com/topic/welfare-state.

institutions even though they bear the responsibility of redistributing wealth and pushing forward legal framework. However, the provision of welfare is not necessarily bound to the administrative guidance of the government, but it can be delivered by the private sector too (i.e., voluntary organizations, companies, religious organizations).[6] Thus, when assessing the impact of AI, it is worth keeping in mind the context and definition of a welfare state to understand the involved parties in the regulatory process of the AI systems.

The international regulatory frameworks are still novel when it comes to AI. In 2019, UNESCO developed a set of AI ethics recommendations to promote the responsible development and deployment of AI systems.[7] The recommendations take into consideration the importance of respecting human rights, promoting human dignity, and benefiting society. As part of the recommendations, an ethical and legal framework for AI systems needs to be created by the states taking into consideration the importance of transparency, accountability, and inclusivity.

However, UNESCO's recommendations are criticized for lacking specific guidelines on how to implement the recommended principles in practice. In addition, recommendations do not address the broader societal and economic implications of AI. Nonetheless, the recommendations provide a valuable starting point for developing ethical frameworks to govern the use of AI in social welfare decision-making and emphasize the importance of ensuring that AI systems are developed and used to respect human rights and promote social welfare. Further work is needed to translate the general principles outlined in the

[6] David Garland, "The Welfare State: A Fundamental Dimension of Modern Government," *European Journal of Sociology / Archives Européennes de Sociologie / Europäisches Archiv Für Soziologie* 55, no. 3 (2014): 327–64.
[7] UNESCO, "193 Countries Adopt First-Ever Global Agreement on the Ethics of Artificial Intelligence," UN News, November 25, 2021, https://news.un.org/en/story/2021/11/1106612.

recommendations into concrete guidelines and best practices for social welfare decision-making by taking into consideration the context, culture, and geography of the new policies.

On the other hand, the EU AI HLEG composed of 52 well-known experts in the field who drafted guidelines on the Trustworthiness of AI brings more practical recommendations for regulating AI. Trustworthiness is divided into three categories. Firstly, AI systems should be lawful throughout their life cycle. Secondly, it should be ethical and thirdly it should be robust in terms of its technical and social implementation following the principles of do no harm.[8] Although it is a new introduction, the EU's efforts in regulating AI are considered progressive, as it follows a similar pattern of introducing the General Data Protection Regulation (GDPR).

This is an important step because, with the rise of automated technologies, the role of governments is also at stake as these technologies are being developed by big corporations like Apple, Google, Microsoft, Meta, and others. According to Harvard Review Writer Christina Pazzannese, there is a lack of oversight by the United States government of private companies using artificial intelligence software to determine health and medicine, employment, creditworthiness, and even criminal justice without having to explain how they are ensuring that programs are not influenced by structural biases, either consciously or unconsciously.[9] In her article, she addresses the governance and ethical aspects of using AI by bringing experts to the field. She argues that the Federal government lacks the expertise and necessary human capital to respond to emerging ethical threats caused by AI leaving the companies to self-police. Thus, such developments create both ethical and socio-economic concerns as the government is responsible for the well-being of people, not corporations. Even though the article does not

[8] HLEG, "Ethics Guidelines for Trustworthy AI" (EU Commission, 2019), https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419.
[9] Christina Pazzanese Harvard Staff Writer, "Ethical Concerns Mount as AI Takes Bigger Decision-Making Role," *Harvard Gazette* (blog), October 26, 2020, https://news.harvard.edu/gazette/story/2020/10/ethical-concerns-mount-as-ai-takes-bigger-decision-making-role/.

address the use of AI in social welfare decision-making, it certainly offers a perspective on understanding ethical questions.

With a similar mindset, Jo Ann Oravec (2019) inserts contributions on the use of AI in social welfare and how every technological advancement affected human decision-making processes regarding social welfare.[10] The author discusses two main terminologies that are essential to consider such as '*sociopolitical cybernetics*' which stresses the loss of human oversight over decision-making processes which might lead to potential disruption in the political system due to AI's domination in creation, maintaining, and changing of social systems. Another term that is often highlighted is the overstatement and hyperbole notion of AI which creates false hopes that eventually create anxiety among those who are concerned that they will be affected by AI. While the author emphasizes primarily the danger of AI in decision-making processes, which is not the same as ethical danger, the latter can be a side effect of the former.

On the other hand, Carney (2021) discusses a more specific case of deploying AI in social welfare in Australia where conclusions regarding AI have been mixed.[11] The paper focuses primarily on the negative impacts of AI on the Australian public due to a rushed and poor understanding of the context that causes a lack of trust among the public. The author argues that before the deployment of AI in social welfare decision making one needs to understand the functional impacts of AI whether it has a supportive role to humans in decision-making, or it can replace humans, and last, if it can disrupt if it is trained on biased data. The author presents his arguments based on the case of Centrelink Online Compliance Initiative (OCI), often known as robodebt, which demonstrates the risks of attempting to install AI systems

[10] Jo Ann Oravec, "Artificial Intelligence, Automation, and Social Welfare: Some Ethical and Historical Perspectives on Technological Overstatement and Hyperbole," *Ethics and Social Welfare* 13, no. 1 (January 2, 2019): 18–32, https://doi.org/10.1080/17496535.2018.1512142.
[11] Terry Carney, "Artificial Intelligence in Welfare: Striking the Vulnerability Balance?," *Monash University Law Review. Monash University. Faculty of Law* 46 (March 16, 2021): Advance, https://doi.org/10.26180/13370369.v1.

with loftier ambitions than current technology and data quality allow. Despite the potential transformative influence of AI in social security, the OCI system proved incapable of turning Australian Tax Office information into the fortnightly earnings data required by social security rules, resulting in erroneous debts and major human and financial expenses.[12]

The use of AI in Australia further has been extended by Whelan and James (2022) [13] who tackle the issues concerning a nonbinding ethical framework on AI drafted by the Department of Industry, Innovation and Science and Data61 (Australian National Science Agency) that includes eight principles focusing on social and environmental well-being of humans, fairness, protection of privacy, safety, transparency, human-centered and accountability. Authors note that when it comes to the implementation of large national AI projects it has been challenging to tick all the principles as it is a complex subject. Yet, these incremental developments at least offer a vision and open a dialogue on preventing the negative impact of AI in decision-making procedures.

1.5.1 Disinformation

With the development of AI applications such as Dall-E, ChatGPT, Speech recognition AI tools, and others it has become easy for malicious actors to use such technology on spreading misinformation. According to Bontridder and Poullet (2021), the rise of disinformation especially during the COVID-19 pandemic created several barriers for governments to fight fake news, conspiracy theories, and manipulative content online.[14] It is also referred to as infodemic by the WHO meaning that there is an excess amount of information to deal with.[15]

[12] Ibid
[13] Alexandra James and Andrew Whelan, "'Ethical' Artificial Intelligence in the Welfare State: Discourse and Discrepancy in Australian Social Services," *Critical Social Policy* 42, no. 1 (February 1, 2022): 22–42, https://doi.org/10.1177/0261018320985463.
[14] Noémi Bontridder and Yves Poullet, "The Role of Artificial Intelligence in Disinformation," *Data & Policy* 3 (ed 2021): e32, https://doi.org/10.1017/dap.2021.20.
[15] "Infodemic," accessed May 5, 2023, https://www.who.int/health-topics/infodemic#tab=tab_1.

While used unethically, AI-powered disinformation campaigns not only can dilute the public but can also disturb the social welfare system by limiting certain types of individuals (minority groups, race, gender, social status) from receiving benefits.

The credit for these new challenges goes to the deep fakes that are powered by Generative Adversarial Networks (GAN).[16] In its basic form, it is a self-learning network that can analyze X amount of content and generate a similar Y content. The spread of these developments can be hazardous in communities where echo chambers are prevalent, and people receive information from restricted sources that limit their critical thinking abilities. However, Bontridder and Poulet argue that if the content that is powered by GANs will be labeled at the moment of its creation it will be easier for the state and public to detect fake content. Another solution to fight deepfakes can be storing individuals' data such as location, demographics, and movement, but this approach would violate human rights causing more harm than help.

The roots of widespread deepfakes come from social media. Yet, despite the negative implications, some authors argue that the use of AI in social media does not raise concerns among the public. In this regard, Isbanner et al. (2022) who examined the use of AI in health services found that respondents to the survey in Australia had negative views when AI was deployed in healthcare in comparison to its general use such as in social media.[17] One of the conclusions in this regard is that many perceive healthcare as a higher-risk area and have less trust in the ethical use of technology in this sphere.

[16] Bontridder and Poullet, "The Role of Artificial Intelligence in Disinformation."
[17] Sebastian Isbanner et al., "The Adoption of Artificial Intelligence in Health Care and Social Services in Australia: Findings From a Methodologically Innovative National Survey of Values and Attitudes (the AVA-AI Study)," *Journal of Medical Internet Research* 24, no. 8 (August 22, 2022): e37611, https://doi.org/10.2196/37611.

Furthermore, one needs to question the rapid growth of these technologies and ask why there has been a mass use of AI if they have negative outcomes. Partly it is due to automation bias.[18] The use of AI in decision-making came with the idea to fix or aid humans to solve social issues such as racism, predicting crimes, and seeking justice. In the United States, for example, predictive policing technologies were promoted "as one answer to racially discriminatory policing, offering a seemingly race-neutral, 'objective' justification for police targeting of poor communities".[19]Numerous other countries such as the United Kingdom, the Netherlands, and Germany have followed such actions, with police forces relying on predictive AI technologies. Similarly, worries about human bias and discrimination prompted the implementation of predictive risk assessment systems in criminal justice.[20] Such systems themselves are identified as causes of bias. One of the recent cases of automation bias occurred in the Netherlands in 2022 when the Dutch Tax and Customs Administration admitted that there was institutional racism in the distribution of childcare benefits due to automation.[21] To create risk profiles the Tax office used algorithms that excluded and labeled foreign-sounding names and dual nationalities as possible fraud. The results were immense as twenty thousand low- and middle-income families fell victim to false accusations of fraud. This mistake cost the government 500 million euros to fix.

1.5.2 Involvement of Different Stakeholders

[18] Saar Alon-Barkat and Madalina Busuioc, "Human-AI Interactions in Public Sector Decision-Making: 'Automation Bias' and 'Selective Adherence' to Algorithmic Advice," *Journal of Public Administration Research and Theory*, February 8, 2022, muac007, https://doi.org/10.1093/jopart/muac007.
[19] Ferguson Andrew, "The Rise of Big Data Policing," *TechCrunch* (blog), October 22, 2017, https://techcrunch.com/2017/10/22/the-rise-of-big-data-policing/.
[20] Ellora Thadaney Israni, "Opinion | When an Algorithm Helps Send You to Prison," *The New York Times*, October 26, 2017, sec. Opinion, https://www.nytimes.com/2017/10/26/opinion/algorithm-compas-sentencing-bias.html.
[21] "Parliamentary Question | The Dutch Childcare Benefit Scandal, Institutional Racism and Algorithms | O-000028/2022 | European Parliament," accessed April 16, 2023, https://www.europarl.europa.eu/doceo/document/O-9-2022-000028_EN.html.

One of the challenges with the aforementioned incidents is due to the lack of appropriate up-to-date policies adopted by the majority of stakeholders regarding the implication of AI in decision-making in social welfare systems. The lack of policies also might be due to the trust culture of specific communities across the world. The research done by Johann Caro-Burnett and Shinji Kaneko (2022) states that trust in the use of AI also differs from place to place and the political atmosphere in a particular context. According to their experiment, the authors argue that in communities where people have a lack of trust in their government will be less likely to evaluate ethical decisions made by AI on its deployment.[22] Although, this research is not robust but theoretical as one can observe from the case of the Netherlands which is a fairly democratic state.

The solution or preventive measures to the failures of AI could be solved by using a whole-of-society[23] approach that allows the participation of all layers of society to critically evaluate AI applications before their deployment to maintain the integrity of political institutions. The whole-of-society approach acknowledges the role of different actors and joint obligations to maintain public integrity and impact public decisions. Not only companies but also individuals have the option of either ignoring the consequences of corruption or actively contributing to the enhancement of public integrity. The OECD Public Integrity Recommendation emphasizes the significance of including all important stakeholders in the formulation, regular updating, and implementation of the public integrity system that can be well applicable while using AI in social welfare.

Promoting public integrity for businesses entails adhering to environmental and human rights norms, paying their fair share of taxes, not giving bribes, and aligning lobbying operations with

[22] Johann Caro-Burnett and Shinji Kaneko, "Is Society Ready for AI Ethical Decision Making? Lessons from a Study on Autonomous Cars," *Journal of Behavioral and Experimental Economics* 98 (June 1, 2022): 101881, https://doi.org/10.1016/j.socec.2022.101881.
[23] OECD, *OECD Public Integrity Handbook* (OECD, 2020), https://doi.org/10.1787/ac8ed8e8-en.

long-term sustainability goals.[24] While initiating policies on social welfare distribution with the help of AI systems these are all essential components of a whole-of-society approach to promoting public integrity. Lastly, it is worth noting that the role of INGOs[25] in AI governance is vital. As Yuval Noah Harari puts it well "what nukes are to the physical world, AI is to virtual and symbolic worlds..." although AI is not limited to the virtual world as we see from the rise of robots.

Nevertheless, like the security measures to prevent nuclear fallout there is a need for an international framework on AI governance. In this regard, this paper will use the case of the EU AI Act proposal as the latest policy development aimed at regulating AI. The EU AI Act proposal is unique in the sense that it adopts a horizontal risk-based approach that applies to all sectors. In comparison to the regulatory system in the US which is rather sector-specific (i.e., the AI developed by the Department of Homeland Security only deals the security issues) the EU AI Act tries to capture all sectors.[26] Thus, this Act will serve both as a literature and a case study to assess the use of AI in social welfare.

In addition, the EU AI Act introduces a framework to install a deliberative process, meaning that the laws on AI ethics and decision-making are legitimate only when they are the product of a public, inclusive dialogue among equals. This opens a space for public-private partnerships where private companies like Google or Microsoft are closely cooperating and co-creating AI projects. In this regard, Russian engineer Cyril Velikanov proposed a concept of "Mass Online Deliberation" which combines human judgment and the technical capacity of AI for storing

[24] Ibid
[25] Eugenio V. Garcia, "Multilateralism and Artificial Intelligence: What Role for the United Nations?," SSRN Scholarly Paper (Rochester, NY, December 12, 2020), https://doi.org/10.2139/ssrn.3779866.
[26] *The European Approach to Regulating Artificial Intelligence*, 2022, https://www.youtube.com/watch?v=BBmq4T_550U.

ideas and proposals into a single appealing vision of deliberative democracy for people.[27] Authors argue that AI can help to scale deliberative democracy by gathering people online in raising certain public issues or discussion of mini clusters of issues which would require fewer resources than organizing the same events in physical space. However, here again, one needs to ensure transparency and ethical use of personal data among those who engage in deliberation.

Chapter I EU AI Act and its possible guidelines on Social Welfare

The EU AI Act proposal is still under review however, general provisions and direction of written regulations in the Act provide a ground to build on the possible future policy outcomes regarding social welfare as it is already reviewed by the board of experts and several hundred organizations across the globe.

Before diving into the implication of the AI Act in social welfare, it is useful to break down the particularities of the document regarding the content. The AI Act consists of twelve Titles (or chapters) each considering a specific theme with general objectives which can be divided into four sections[28]:

1. Ensuring safe use of AI applications in the EU market with respect to EU values and laws.
2. Clarifying legal frameworks to develop investment and innovation in the AI domain.
3. Improving the governance structures along with an enforcement mechanism of existing laws and building safety requirements for AI applications.

[27] "Hélène Landemore: Can AI Bring Deliberation to the Masses?," Stanford HAI, accessed October 28, 2022, https://hai.stanford.edu/events/helene-landemore-can-ai-bring-deliberation-masses.
[28] "Institutional Context," *The Artificial Intelligence Act* (blog), December 16, 2022, https://artificialintelligenceact.eu/context/.

4. Developing a single market for legal and safe use of AI applications and preventing market divisions.

These objectives are proposed using a horizontal regulatory approach regarding the use of AI that applies to all industries with a minimum required consideration of the risks and challenges related to AI. The benefit of using this approach helps to hinder the overregulation of technological progress and disproportional distribution of costs of developing AI systems and harmonization of the market.[29]

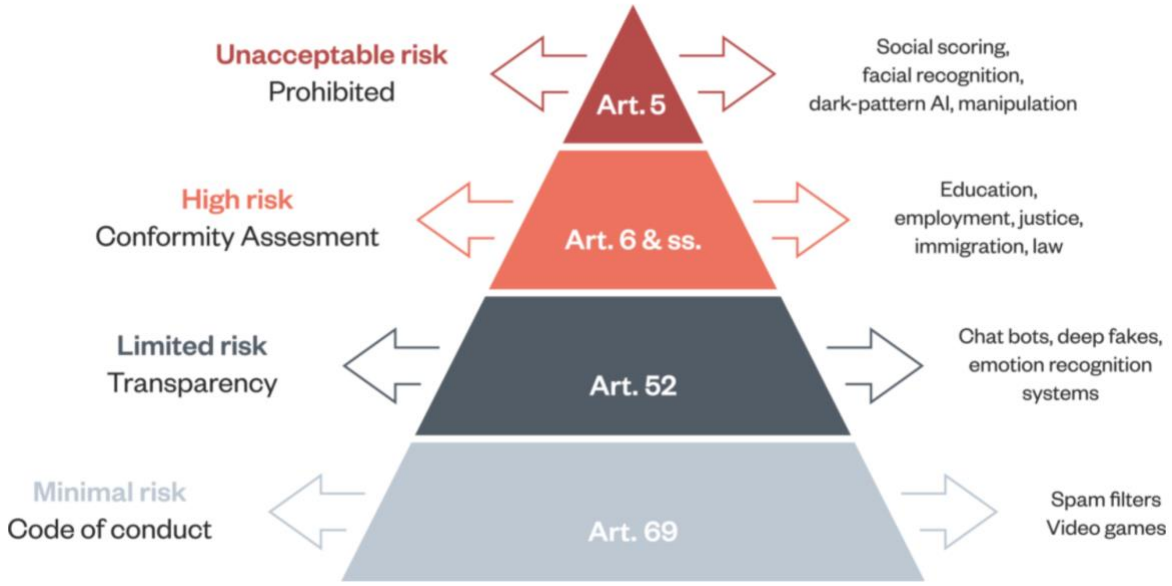
In addition, the regulatory framework is built on already existing laws and consistent with the EU Charter of Fundamental Rights along with legislatures on data protection (General Data Protection Regulation) – GDPR (Regulation (EU) 2016/679), non-discrimination, and gender equality.[30] This approach makes the proposal future-proof to the new emerging technologies and complies with principal requirements. Also, the horizontal application of AI systems is equally applied to EU member states in a binding order to make a single regulatory framework.

Like other regulatory frameworks, when it comes to the content and the goal of the AI Act, the priority leans towards addressing the risks caused by AI. The risks in the AI Act are divided into four categories[31]:

- unacceptable risk (Title II),
- high risk (Title III),

[29] EU Commission, "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," April 4, 2021, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206.
[30] Frederik J. Zuiderveen Borgesius, "Strengthening Legal Protection against Discrimination by Algorithms and Artificial Intelligence," *The International Journal of Human Rights* 24, no. 10 (November 25, 2020): 1572–93, https://doi.org/10.1080/13642987.2020.1743976.
[31] EU Commission, "The AI Act Proposal," February 10, 2021, https://artificialintelligenceact.eu/the-act/.

- restricted risk (also known as systems that possess the danger of manipulation) (Title IV), and
- risks of all other AI applications (Title IV).



Source: Ethical Intelligence Blog[32]

1.6 The indirect impact of AI on Social Welfare Decision-Making

While the categorization of risks does not require much effort, the assessment of AI applications and judgment over its potential risks can be challenging especially when they are not always obvious. On the one hand, risk categories help to understand the typology of deployed AI applications. For instance, the section on the protection of Fundamental rights[33] of AI Act proposes to categorize AI systems as high risk if they are used to screen job applications, recruitment processes, monitor, or make decisions for promotion. On the other hand, these categories remain vague and open to changing views which ultimately relies on the court's interpretation of the categorization taking into account the context. In the scenarios

[32] "The Impending EU AI Act - What You Need to Know — Ethical Intelligence," EI, accessed May 12, 2023, https://www.ethicalintelligence.co/blog/the-impending-eu-ai-act-what-you-need-to-know.
[33] EU Commission, "The AI Act Proposal," page 26. February 10, 2021, https://artificialintelligenceact.eu/the-act/.

where AI is responsible to assess the creditworthiness of an individual, it can be even riskier if used improperly as many people depend on credit to have access to basic needs such as electricity, housing, and telecommunications.[34]

Nonetheless, until the review of courts AI and Automated Decision Making (ADM) systems that are wrongfully assigned high-risk instead of unacceptable risk, can bring harm to society and especially to vulnerable groups (people of color, women, children, migrants) which is often the case due to data deficiency. While in use, these systems could result in indirect negative consequences. To illustrate, if high-risk AI systems are deployed by the police to profile layers of society or for surveillance reasons (i.e., facial recognition programs) then those citizens who are unjustly accused of not following public order might be jailed until the decision of the court. During the period of detention, those individuals could be ripped of rights to access social welfare. Further, it might create a stigma and negative consequences to receive access to social services.[35]

In addition, the surveillance might work in a vice-versa format, where individuals are policed based on their social background based on their access to social welfare, previous criminal record, and government subsidies. One such case was revealed during the investigation of Big Brother Watch, an organization for rights in the UK where one of the postcode indicators used by Durham Police as a predictive policing method in their Harm Assessment Risk Tool (HART). Police used data that was originally designed as a commercially available marketing product called 'Mosaic.' This product, created and sold by Experian, a global data broker, profiles and categorizes areas (postcodes) in the UK into socio-geodemographic profiles. Experian claims that this product provides a detailed picture of today's UK consumer by using

[34] Stanford Law School, "EU Artificial Intelligence Act: The European Approach to AI," Stanford Law School, accessed May 12, 2023, https://law.stanford.edu/publications/eu-artificial-intelligence-act-the-european-approach-to-ai/.
[35] Ensure rights and redress for people impacted by AI systems, EDRI

information such as census data, land registry data, exam results, welfare and benefits data, and family/personal names linked to ethnicity, among others. This product includes more than 850 million data points.[36] As a result, this gives police immense power on making predictive decisions based on the postal code which is significantly biased and inaccurate.

The AI Act calls for classifying such AI systems as high-risk if they are used by authorities to decide on the denial, reduction, revocation, or reclamation of benefits and services, as such decisions can substantially affect people's livelihood and may violate their fundamental rights, including the right to social protection, non-discrimination, human dignity, and an effective remedy. This follows by the concerns raised by many civil society voices who urged for the prohibition of practices that draw conclusions about individuals' eligibility for social welfare based on the assumption that they may be more likely to commit criminal offenses.[37] To prevent such incidences, the AI Act calls for the balanced and transparent use of AI technologies to ensure the fundamental rights of the individual.[38] This in return enables companies that are using AI systems to lay out risks and provide further information to build trust among their customers and also for public institutions the transparent understanding of using AI will bring public trust.

However, the trust is subjective and might change and the regulatory mechanism also takes into account that there should be sustainable independent audits, adaptive measures, and certification mechanisms. AI Act states that certificates shall be issued by notified institutions that specify in specific fields where the AI will be deployed (i.e., pension system, tax authorities, healthcare, machinery) in accordance with EU laws and language. To ensure their

[36] Ibid
[37] Fair Trials, "Automating Injustice: The Use Of Artificial Intelligence & Automated Decision-Making Systems In Criminal Justice In Europe," 2021, https://www.fairtrials.org/app/uploads/2021/11/Automating_Injustice.pdf.
[38] EU Commission, "The AI Act Proposal."Page 27

validity and credibility the certificates should not be issued for a period exceeding 5 years. The extension of validity might be subject to re-assessment considering risk categories.[39]

Furthermore, Articles 43 and 47 of the Act foresee the conformity of the high-risk applications in use. Member states are able to raise objections on high-risk AI applications in use if the objection is justified by the Commission these applications should be taken down from the market. Special permission can be granted for high-risk AI systems in certain situations, such as public security or protecting people's lives and the environment. If no objections are raised within 15 days, the authorization is considered valid. However, if objections are raised or the authorization is deemed unjustified, consultations and a decision-making process involving relevant parties will take place.[40]

In addition, when it comes to the providers of high-risk AI systems Article 52 of the Act presents the following guidelines:

- Individuals must be informed when they are interacting with an AI system unless it is already obvious.
- However, this does not apply to AI systems used for law enforcement purposes, unless they are publicly accessible for reporting crimes.
- Users of emotion recognition or biometric categorization systems must also inform individuals about the operation of the system, except for systems used in law enforcement.
- If an AI system is deemed to pose a risk to people's health, safety, or fundamental rights by a market surveillance authority, the operator may be required to take appropriate measures to mitigate the risk, withdraw the system from the market, or recall it.

[39] EU Commission.
[40] EU Commission. Page 64

- The authority will inform the Commission and other Member States about the details of the AI system and the actions taken.[41]

However, the implication of these articles in practice is rather difficult due to the technical capabilities of certain member states and human resources to audit these systems in a timely manner. This in return might negatively impact not only social welfare but also other critical areas of the economy.

Chapter II: Risk of AI in Social Welfare

This chapter focuses on the future implications of AI systems through the lenses of the AI Act in social welfare systems. It assesses the considered safety measures for the safe deployment of AI systems in the EU market. Mainly the debate within this chapter will shed light on the progress versus safety measures or in other words regulatory barriers to the deployment of AI systems in decision-making processes.

One of the downsides of the safety measures used in the AI Act is the risk-based approach to determining the acceptability of the AI systems entering the market on an *ex-ante* (before the event) basis. There are several problems involved in using a risk-based approach and civil society along with researchers has been vocal about not using a risk-based approach while protecting human rights. According to Access Now, digital rights organizations, a risk-based approach is beneficial in specific technical fields where companies need to assess the risk of their operation. The outcome of this approach put by the EU leads to the situation where the evaluation of operational risks of companies is equal to people's basic rights which creates an imbalance between human rights and the profit-oriented nature of companies.[42]

[41] EU Commission. Page 69
[42] Fanny Hidvegi Massé Daniel Leufer, Estelle, "The EU Should Regulate AI on the Basis of Rights, Not Risks," *Access Now* (blog), February 17, 2021, https://www.accessnow.org/eu-regulation-ai-risk-based-approach/.

Doing so breaches fundamental human rights for two reasons. Firstly, the prior assessment depends on the context in which the AI system is being used. Here the horizontal approach of the EU would be challenged because EU states are not the same and the perception of risks differs from culture to culture. Secondly, the companies would have an incentive to downgrade the associated risks involved in their AI applications. As a result, low-risk AI applications will be left with almost no oversight and safety regulations. For these reasons, using a risk-based approach could potentially harm human rights which is also a fundamental to consider in social welfare decision-making processes.

To add on risks, since several European countries have already implemented AI in their social welfare systems, it is imperative to ensure that human rights are prioritized over profit or cost-cutting. The government's rationale for adopting AI in social welfare which revolves around cost-cutting measures efficiency is largely benefitting the government but not the people. For instance, in 2018 the Public Employment Service in Austria (Arbeitsmarktservice, AMS) started using an algorithmic profiling system to profile job seekers. The profiling is divided into three distinct categories such as[43]:

- Category A – applicable to those with high job prospects in the short term
- Category B – applicable to those with mediocre job prospects which represents a larger portion of job seekers.
- Category C – applicable to those with low skills and low prospects

In accordance, with the category, individuals receive different treatment and assistance. For the government on the other hand this categorization helps to hand-pick programs or design policies targeting category B since it represents most of the job seekers and leaves the category

[43] Doris Allhutter et al., "Algorithmic Profiling of Job Seekers in Austria: How Austerity Politics Are Made Effective," *Frontiers in Big Data* 3 (2020), https://www.frontiersin.org/articles/10.3389/fdata.2020.00005.

A and C out of the loop as they are not cost-effective. The justification behind this system of categorization is due to the numbers that anyone who falls under Category A has a probability of finding a job within 3 to 7 months and there is no need for huge investment in them. While Category C represents less than 25% of the whole job seeker group which might take more effort for the government to compensate thus it chooses to leave it for external organizations to deal with.[44] As a result, vulnerable groups in Category C such as women, the disabled, youth, or migrants might receive only a little welfare assistance. Precisely, the notion of classifying people with a migration background is in itself problematic as it creates s stigma around these individuals as some sort of secondary priority people. As a result, an overwhelming majority of individuals with a migration background are put into Category C.[45] According to the president of AMS, these actions do not violate discrimination laws because people are not categorized according to their ethnic descent. Moreover, AMS does not clarify the concept of migration descent, which might likely be judged as an ethnic descent.[46] In addition, the algorithm used by AMS builds normative data both at the individual and group level along with social status. This presents a clear picture for the government not only target groups but also individuals and communities to design different labor policies that ultimately will entail discriminatory results. According to the AI Act, such algorithmic classification falls under high-risk (Title III)[47] applications and will need to be scrutinized and tested before they will be used in social welfare programs.

1.7 Preventive Measures

One of the foreseen measures in the AI Act in this regard introduced by the High-Level Expert Group on Artificial Intelligence (HLEG) is the introduction of methods like Sandbox.

[44] Allhutter et al.
[45] Allhutter et al.
[46] Allhutter et al.
[47] EU Commission, "The AI Act Proposal."

Sandboxes are introduced to ease the deployment of AI in controlled environments to have relatively safer applications in real practices.[48] The method of sandboxes is especially helpful for small and medium-sized Enterprises (SMEs) and start-ups that are looking to enter the AI market but lack in resources to test and build comprehensive safety measures.

The controlled environment of Sandboxes allows using of separate protected data isolated from other environments to test the AI system along with controlled authorization of eligible parties to have access to the test data. In ideal conditions, this in return will ensure the progress of innovation in the sector.[49]

However, when it comes to the application of AI systems in social welfare associated risks and the scale become significantly bigger. Again, going back to the case of the AMS algorithmic application, one can understand that classification of citizens is part of the agreed agenda within the government and cuts costs. Yet initially the application was introduced to increase the efficiency of providing jobs for job seekers. This mere practice can be further practiced by governments with nationalistic agendas and the AI regulation that will be in place cannot address these issues promptly. As a result, two concerns or questions emerge regarding testing high-risk AI applications, particularly in social welfare:

1. Fairness and Bias: During the testing period of high-risk applications in social welfare one potential threat is unfair outcomes How one can ensure that the testing process adequately fits the needs of a particular EU country or context given different populations and cultures?

[48] Mauritz Kop, "EU Artificial Intelligence Act: The European Approach to AI - Journal Article - Stanford Law School," accessed May 13, 2023, https://law.stanford.edu/publications/eu-artificial-intelligence-act-the-european-approach-to-ai/.
[49] "Artificial Intelligence Act and Regulatory Sandboxes | Think Tank | European Parliament," accessed May 29, 2023, https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2022)733544.

2. Accountability and Transparency: Many already existing AI-powered tools used by the governments in Netherlands or Austria are not transparent in their operations. How those algorithms are tested and received is essential to build accountability among the public.

Title V of the AI Act states that each state-appropriate institution should create sandboxes to address the aforementioned two points, but ultimately the success of these testing tools will rely on the honesty and accountability of assigned institutions and socio-political agents to determine future directions of using AI in social welfare systems.

Chapter III: Discussions & Conclusion

In the previous chapters, general provisions of the EU Artificial Intelligence Act proposals were presented along with case studies involving algorithms in the social welfare sector. In this chapter, however, we shall examine the shortcomings of the AI Act and summarize the main points that have been discussed regarding the implications of the Act for social welfare.

In a nutshell, the challenges of harmonizing AI and its horizontal harmonizing applications across all sectors can be separated into five distinct sections[50] such as:

- **Complexity and interconnectivity:** Almost all AI systems that are including in the social welfare decision-making process are complex and trained on multiple data sets, codes, and in different demographic contexts. For instance, reflecting on the case of using algorithmic applications at the Austrian Public Employment Services (AMS) agency one can realize such complexity and interconnectedness of diverse sets of systems within the categorization of job seekers, using concepts and laws on labor

[50] "Standardizing AI - The Case of the European Commission's Proposal for an Artificial Intelligence Act by Martin Ebers :: SSRN," accessed May 20, 2023, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3900378.

regulation, application of differing standards, and others. As a result, this complexity makes it challenging to monitor high-risk and unacceptable-risk AI applications to make a case for breaches of the law.

- **Correlation instead of causation:** The majority of AI systems are trained in a way to identify statistical correlations and patterns instead of reflecting on causation among available datasets. To ensure fair and unbiased results the AI systems must be trained using multiple diverse datasets to prevent adverse effects on reinforcing existing biases or human-made mistakes.
- **Adaptation to the context:** The AI systems in comparison to other technologies are unique in the sense that it is hard to pace their development and its results. Given their ability to correlate and learn from past actions AI systems quickly adapt to the changing environment or posed questions in applications like ChatGPT[51] and others. The probabilistic features of such systems make it challenging to propose a standard quality across all sectors as these systems often hallucinate and cannot give 100% right answers. One difficulty in applying horizontal implications of AI regulations stems from the fact that while traditionally there were agreed-upon axioms with mathematically verifiable measures, such as the value of P being approximately equal to 3.14..., AI systems pose challenges in establishing universally verifiable metrics. For example, AI applications can nominate candidate X for the director position at Y Company using (trained) Dataset A but discard the candidacy of X if trained using Dataset B. So, ultimately AI's adaption to the context is highly dependent on the use of data which plays an essential reason for future outcomes.
- **Autonomous behavior**: Some AI systems under the high and unacceptable risk categories might propose decisions without human oversight which might violate

[51] "Introducing ChatGPT," accessed May 25, 2023, https://openai.com/blog/chatgpt.

human rights and violate safety rules. These types of AI applications are already in use in multiple sectors including in the hiring or screening processes of job applicants. Despite the indicated parameters by humans while making decisions, AI still can-do mistakes and reject applications.

- **Transparency:** Lastly, the lack of transparent processes behind these systems leaves less hope for their long-term benefits. Even in the scenario where companies can reveal the used code and regulations used behind these will be open to the public, it will be hardly possible for ordinary people to understand the basic Terms and Conditions of these systems. The consequences of this would later bring more issues to protect human rights.

1.8 The Use of AI Software

Further shortcomings that are essential to discuss is the use of AI-powered software in specific contexts by public authorities during decision-making. According to the AI Act proposal Articles 3 (2), and Article 28 (1) on providers, distributors, importers, and users, one can realize an unbalanced scheme. For instance, if AMS services develop their AI-powered software for categorizing job seekers, then the agency, under the proposed AI Act would be a "provider"[52] of a "high-risk" AI system subject to Article 6 of the Act. So, in this case, AMS is required to comply with the risk prevention measures of their software which would include providing information about the periods of using the software and establishing processes that assign their employees to have control over the mistakes of the software.

However, if AMS would buy software from a tech company to help with the categorization of job seekers, then they would be considered as "users" instead of a provider.[53] The problem then here emerges from the fact that the AMS is no longer responsible for disclosing information

[52] EU Commission, "The AI Act Proposal," Articles 3 (2) & 28 (1) February 10, 2021, https://artificialintelligenceact.eu/the-act/.
[53] Ibid Article 3

on using the technology, giving them a wide range of discretion to outsource their work to third-party software providers.[54] Oftentimes, policymakers miss addressing other such issues with an assumption that providers will bear the responsibility and follow the regulations by signing the conformity agreement before launching their product in the EU market. However, in the practice of providing social services ultimately it is the responsibility of the public authority to ensure that the fundamental rights of people are prioritized. In addition, if the discretion is left in the hands of companies, then they might downgrade the risks of their applications to increase their production. Moreover, even on the occasion when companies are left to self-regulate their products, it will leave little space for ultimate users such as civil society, and ordinary people who are affected by the decisions made by AI to participate in the development stages of the software. As a result of these gaps, the AI Act fails to offer explicit guidance to individuals directly impacted by AI software decisions on the proper procedures for lodging complaints and taking action against providers in cases where they may have been erroneously denied social benefits.[55]

On the other hand, AI Act also downplays the risks and regulatory frameworks when it comes to minimal-risk AI systems. According to the Act providers of minimal risk only bear a limited transparency of information while for high-risk AI systems besides costs involved in the testing period, they must provide high-quality data and documentation along with features to trace the activities of such systems.[56] However, if providers of minimal-risk AI systems wish to ensure that their products are trustworthy then the costs of testing would be as high as for high-risk applications which is not a viable option for early developers.

[54] "How the EU's Flawed Artificial Intelligence Regulation Endangers the Social Safety Net: Questions and Answers | Human Rights Watch," accessed May 25, 2023, https://www.hrw.org/news/2021/11/10/how-eus-flawed-artificial-intelligence-regulation-endangers-social-safety-net#_ftn65.
[55] "How the EU's Flawed Artificial Intelligence Regulation Endangers the Social Safety Net: Questions and Answers | Human Rights Watch."
[56] EU Commission, "The AI Act Proposal."(page 10)

Essentially, it comes to the point where regulators and public authorities need to take a human rights-based approach rather than a risk-based approach. However, the deployment of AI is largely dependent on its future users and the context in which it will be deployed. Thus, it is not only challenging to determine the risks but also given the rapid pace of AI development it is difficult to have a future-proof regulation.

1.9 Regulation vs Progress

While AI is not going anywhere and is here to stay, the further step would be building a balance between regulation and progress. Since its proposal, the AI Act in the eyes of many developers has been perceived as a step towards the regress of technological advancements, thus creating two fronts. As a result, on one side, proponents are advocating for regulation and a more cautious approach toward the development of AI. On the other side, there are those who hold a neoliberal perspective, arguing for minimal intervention and relying on market forces and individual regulation.

The often-cited pro argument for the regulation is that in the near future, the regulation of AI might slow down the pace of progress, especially in the startup world but at the same time, in the long run, the regulation is more beneficial for the end users of the product than for companies who are oriented on making profits through fast and destructive measures. Now the question on the table is to what extent the introduction of sandboxes in the AI Act will be enforced to prevent potential harms and assign appropriate risk measures so that it brings less burden on the progress of growing companies.

Conclusion

In a recent interview, Google's CEO Sundar Pichai acknowledged that "they do not have all the answers yet…" referring to the use of AI and the way it interacts.[57] This message is both alarming and also a call for the introduction of robust regulatory systems that can prevent potential harms of AI to protect human rights and ensure the democratic participation of all layers of society in the design and implementation of AI applications.

The purpose of this research was to understand to what extent the current regulatory systems are responding to the risks caused by AI. Despite, novel steps with pitfalls in addressing specificities of risks in social welfare and beyond, the EU AI Act proposal serves at least as a stepping stone at the international level to prevent possible disruptions among all sectors of the economy.

Throughout the paper and presented case studies on the use of AI in social welfare it is clear that AI can have negative consequences (bias, unfairness) and potentially infringe human rights. It became apparent that there is a need for better binding regulatory systems and the EU AI Act proposal's horizontal approach brings appropriate measures to address these issues. The approach helps to standardize the AI sector under a single regulatory umbrella across the union to harmonize the market. Although, as it was presented in the analysis of the Act in social welfare, the risk-based approach to assessing high-risk applications before their deployment can bring harm to individuals and enable companies to downgrade the risk level of their product. It was also noted that if companies are left to regulate themselves civil society or other

[57] *Google CEO Shares His Concerns about AI | 60 Minutes*, 2023, https://www.youtube.com/watch?v=MJs-1QxWCbI.

independent parties will not be part of the solution to provide safe products and by doing so disrupt the whole of society approach to decision-making processes.

The paper also highlighted a possible solution in this regard including the introduction of Sandboxes by member states where high-risk AI products go through comprehensive testing before their deployment in different sectors. Yet, doing so also brings up the debate between progress and regress. Small companies or start-ups might not have sufficient resources to test their products and bring them to the market and thus leading to a lack of innovation. However, the argument proposed by the experts of the AI Act relies on the theory that in the short run, regulations might undermine progress, especially start-ups, in the long run, it will ensure the safe deployment of AI systems and ultimately benefits users.

All in all, the deployment and risks of AI systems in social welfare might be divided into near, medium, and future. In the near future or currently one can observe the risks of AI in spreading misinformation and engagement in the political spectrum of left and right. In the medium (3-4) years one can expect to see mass disruptions in the labor market due to the deployment of AI where people again will be heavily reliant on the government in terms of financial and social assistance to back up the costs of unemployment. And lastly, if AI is not regulated and if it will continue accelerating at the current pace, AI applications will become competitive with humans and cause even bigger disruptions in critical spheres that go beyond the focus of current research. While this paper covered the AI Act proposal and its possible implications in social welfare, there are still gaps to address quantifiable risks such as the financial costs of misusing AI in social welfare and benefits in sectors such as critical spheres like healthcare and labor market disruptions.

Bibliography

Allhutter, Doris, Florian Cech, Fabian Fischer, Gabriel Grill, and Astrid Mager. "Algorithmic Profiling of Job Seekers in Austria: How Austerity Politics Are Made Effective." *Frontiers in Big Data* 3 (2020). https://www.frontiersin.org/articles/10.3389/fdata.2020.00005.

Alon-Barkat, Saar, and Madalina Busuioc. "Human-AI Interactions in Public Sector Decision-Making: 'Automation Bias' and 'Selective Adherence' to Algorithmic Advice." *Journal of Public Administration Research and Theory*, February 8, 2022, muac007. https://doi.org/10.1093/jopart/muac007.

Andrew, Ferguson. "The Rise of Big Data Policing." *TechCrunch* (blog), October 22, 2017. https://techcrunch.com/2017/10/22/the-rise-of-big-data-policing/.

"Artificial Intelligence Act and Regulatory Sandboxes | Think Tank | European Parliament." Accessed May 29, 2023. https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2022)733544.

Bontridder, Noémi, and Yves Poullet. "The Role of Artificial Intelligence in Disinformation." *Data & Policy* 3 (ed 2021): e32. https://doi.org/10.1017/dap.2021.20.

Carney, Terry. "Artificial Intelligence in Welfare: Striking the Vulnerability Balance?" *Monash University Law Review. Monash University. Faculty of Law* 46 (March 16, 2021): Advance. https://doi.org/10.26180/13370369.v1.

Caro-Burnett, Johann, and Shinji Kaneko. "Is Society Ready for AI Ethical Decision Making? Lessons from a Study on Autonomous Cars." *Journal of Behavioral and Experimental Economics* 98 (June 1, 2022): 101881. https://doi.org/10.1016/j.socec.2022.101881.

EI. "The Impending EU AI Act - What You Need to Know — Ethical Intelligence." Accessed May 12, 2023. https://www.ethicalintelligence.co/blog/the-impending-eu-ai-act-what-you-need-to-know.

EU Commission. "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," April 4, 2021. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206.

———. "The AI Act Proposal," February 10, 2021. https://artificialintelligenceact.eu/the-act/.

Fair Trials. "AUTOMATING INJUSTICE: THE USE OF ARTIFICIAL INTELLIGENCE & AUTOMATED DECISION-MAKING SYSTEMS IN CRIMINAL JUSTICE IN EUROPE," 2021. https://www.fairtrials.org/app/uploads/2021/11/Automating_Injustice.pdf.

Garcia, Eugenio V. "Multilateralism and Artificial Intelligence: What Role for the United Nations?" SSRN Scholarly Paper. Rochester, NY, December 12, 2020. https://doi.org/10.2139/ssrn.3779866.

Garland, David. "The Welfare State: A Fundamental Dimension of Modern Government." *European Journal of Sociology / Archives Européennes de Sociologie / Europäisches Archiv Für Soziologie* 55, no. 3 (2014): 327–64.

Google CEO Shares His Concerns about AI | 60 Minutes, 2023. https://www.youtube.com/watch?v=MJs-1QxWCbI.

HLEG. "Ethics Guidelines for Trustworthy AI." EU Commission, 2019. https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419.

"How the EU's Flawed Artificial Intelligence Regulation Endangers the Social Safety Net: Questions and Answers | Human Rights Watch." Accessed May 25, 2023.

https://www.hrw.org/news/2021/11/10/how-eus-flawed-artificial-intelligence-regulation-endangers-social-safety-net#_ftn65.
"Infodemic." Accessed May 5, 2023. https://www.who.int/health-topics/infodemic#tab=tab_1.
"Introducing ChatGPT." Accessed May 25, 2023. https://openai.com/blog/chatgpt.
Isbanner, Sebastian, Pauline O'Shaughnessy, David Steel, Scarlet Wilcock, and Stacy Carter. "The Adoption of Artificial Intelligence in Health Care and Social Services in Australia: Findings From a Methodologically Innovative National Survey of Values and Attitudes (the AVA-AI Study)." *Journal of Medical Internet Research* 24, no. 8 (August 22, 2022): e37611. https://doi.org/10.2196/37611.
Israni, Ellora Thadaney. "Opinion | When an Algorithm Helps Send You to Prison." *The New York Times*, October 26, 2017, sec. Opinion. https://www.nytimes.com/2017/10/26/opinion/algorithm-compas-sentencing-bias.html.
James, Alexandra, and Andrew Whelan. "'Ethical' Artificial Intelligence in the Welfare State: Discourse and Discrepancy in Australian Social Services." *Critical Social Policy* 42, no. 1 (February 1, 2022): 22–42. https://doi.org/10.1177/0261018320985463.
Massé, Fanny Hidvegi, Daniel Leufer, Estelle. "The EU Should Regulate AI on the Basis of Rights, Not Risks." *Access Now* (blog), February 17, 2021. https://www.accessnow.org/eu-regulation-ai-risk-based-approach/.
Mauritz Kop. "EU Artificial Intelligence Act: The European Approach to AI - Journal Article - Stanford Law School." Accessed May 13, 2023. https://law.stanford.edu/publications/eu-artificial-intelligence-act-the-european-approach-to-ai/.
"More than 4 Billion People Still Lack Any Social Protection, ILO Report Finds." News, September 1, 2021. http://www.ilo.org/global/about-the-ilo/newsroom/news/WCMS_817653/lang--en/index.htm.
OECD. *OECD Public Integrity Handbook*. OECD, 2020. https://doi.org/10.1787/ac8ed8e8-en.
Oravec, Jo Ann. "Artificial Intelligence, Automation, and Social Welfare: Some Ethical and Historical Perspectives on Technological Overstatement and Hyperbole." *Ethics and Social Welfare* 13, no. 1 (January 2, 2019): 18–32. https://doi.org/10.1080/17496535.2018.1512142.
"Parliamentary Question | The Dutch Childcare Benefit Scandal, Institutional Racism, and Algorithms | O-000028/2022 | European Parliament." Accessed April 16, 2023. https://www.europarl.europa.eu/doceo/document/O-9-2022-000028_EN.html.
School, Stanford Law. "EU Artificial Intelligence Act: The European Approach to AI." Stanford Law School. Accessed May 12, 2023. https://law.stanford.edu/publications/eu-artificial-intelligence-act-the-european-approach-to-ai/.
"Standardizing AI - The Case of the European Commission's Proposal for an Artificial Intelligence Act by Martin Ebers :: SSRN." Accessed May 20, 2023. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3900378.
Stanford HAI. "Hélène Landemore: Can AI Bring Deliberation to the Masses?" Accessed October 28, 2022. https://hai.stanford.edu/events/helene-landemore-can-ai-bring-deliberation-masses.
The Artificial Intelligence Act. "Institutional Context," December 16, 2022. https://artificialintelligenceact.eu/context/.

The European Approach to Regulating Artificial Intelligence, 2022. https://www.youtube.com/watch?v=BBmq4T_550U.
"The OECD Artificial Intelligence (AI) Principles." Accessed May 1, 2023. https://oecd.ai/en/ai-principles.
UNESCO. "193 Countries Adopt First-Ever Global Agreement on the Ethics of Artificial Intelligence." UN News, November 25, 2021. https://news.un.org/en/story/2021/11/1106612.
"Welfare State | Britannica," April 18, 2023. https://www.britannica.com/topic/welfare-state.
World Bank. "Poverty." Text/HTML. Accessed April 30, 2023. https://www.worldbank.org/en/topic/poverty/overview.
Writer, Christina Pazzanese Harvard Staff. "Ethical Concerns Mount as AI Takes Bigger Decision-Making Role." *Harvard Gazette* (blog), October 26, 2020. https://news.harvard.edu/gazette/story/2020/10/ethical-concerns-mount-as-ai-takes-bigger-decision-making-role/.
Zuiderveen Borgesius, Frederik J. "Strengthening Legal Protection against Discrimination by Algorithms and Artificial Intelligence." *The International Journal of Human Rights* 24, no. 10 (November 25, 2020): 1572–93. https://doi.org/10.1080/13642987.2020.1743976.