AI ETHICS

Ethics in AI for Peace

By Branka Panic, Founder and Executive Director, AI for Peace

With rapid developments of data science, the greater availability of data and increased computer processing power, algorithms have become part of many spheres of our lives and work, including in peacebuilding and conflict prevention work. The "algo age" brought with it many technical, legal, and ethical questions that exceeded the scope of existing peacebuilding and peacetech ethics frameworks. This module raises awareness about these risks, and advises on how to apply tools to mitigate them by transferring ethical principles into practice to ensure that "AI ethics by design" is a first step and not an afterthought in AI for peace projects' development and implementation.

Lecture Transcript

0:03 Welcome to Ethics in Artificial Intelligence for Peace. My name is Branka Panic. I'm a founder and executive director at AI for Peace, a US based organization working at the intersection of artificial intelligence and peacebuilding. I'm also a co-founder and board member of the Center for Exponential Technologies connecting policy and tech worlds. And I'm currently a visiting scholar at the New York University Center on International Cooperation, researching the data driven approaches to peacebuilding and conflict prevention

0:45 At AI for Peace, our vision is a future in which artificial intelligence benefits peace, security, and sustainable development and where diverse voices influence creation of AI and related technologies. We are preparing peacebuilders with the mindset, knowledge and tools to embrace research, guide and develop human centered artificial intelligence, ensuring the creation of sustainable positive peace. We are trying to help these builders, humanitarians and human rights activists and all aligned with our vision to be well equipped to benefit from data science infrastructures, and techniques such as machine learning or natural language processing. And we also serve as a global open hub for social scientists for AI researchers, developers, policymakers who want to understand and leverage AI potential and impact.

1:56 So we just want to mention that AI for Peace is a sort of continuation of certain processes that started even with the emergence of first ICT tools or general technologies. So this is part of a bigger ecosystem of digital peacebuilding as a broader nexus between the field of peacebuilding and digital technologies. And also the continuation of a Peacetech is a technology that contributes to peace building, with that difference that at AI for Peace, we are not looking only at positive applications of AI, but we are also looking at those risks and often unintended consequences. We are trying to raise awareness about those issues when these builders or other stakeholders are using artificial intelligence.

2:54 So why are we bringing AI now to the to the peacetech? Why AI? Why now? It's almost not needed to be said that AI is not the future anymore, that AI is already here and affecting people across the planet, and AI applications have already started transforming different industries. So we're looking at these trends as well,

Transcribed by https://otter.ai

including the cost of computer processing power that has fallen dramatically. And then the new generation of human, more powerful tools available to different stakeholders, including peacebuilders. Then, of course, the greater availability of data with the cost of cloud access and data storage and processing that has also fallen. And then especially the the capabilities of machine learning and the increased utilization of ML as a method of data analysis that enables computers to learn without any external instructions.

4:05 So why do we bring peace to this field and why now? So most of our founders are coming from the peacebuilding field and we are working on conflict and violence, seeing them as some of the biggest challenges our humanity is facing today. And conflict and violence are interlinked with problems of climate change, poverty, disease, famine, and they're taking lives and causing harm that can be prevented. So what we are seeing as some very worrying trends is that the world is becoming less peaceful. What the Institute for Economics and Peace is showing us and the Global Peace Index, published in this year 2022 is showing us that the average level of global peacefulness deteriorated by 0.3%. And this is actually the 11th deterioration of peacefulness in the last 14 years. With 90 country, 90 countries improving, but 71 deteriorating, and two remaining stable in peacefulness, highlighting that countries tend to deteriorate much faster than they improve.

5:37 We are also looking into unprecedented humanitarian crisis. And only a couple of years ago, when we looked into these numbers, the 82.4 million people were forcibly displaced. Now, UNHCR is reporting that this number has increased significantly. In the first months of this year, more than 100 million individuals were displaced as a result of persecution, conflict, violence, or human rights violations. Another concerning trend is democracy decline. And the present threat to democracy is the product of 16 consecutive years of decline in global freedom where a total of 60 countries suffered declines over the past year and only 25 improved. So 38% of the global population live in so-called not free countries. And this is the highest proportion since 1997, where only 20% lived in so-called free countries.

6:57 So with all of this, when we started our work, we looked into these connections between technology and peace. Seeing that peace needs allies urgently today, we have a lot of problems and challenges to face. But what we saw is that actually military uses of AI are multiple and advanced. The AI is being used in many fields such as battlefield healthcare, for combat simulation and training, for target recognition. Most of the people are probably most familiar with the so-called autonomous weapons systems or killer robots, then the AI is being used for threat monitoring and situational awareness. But what we saw is that the applications of AI exploring and utilizing AI in peacebuilding is very limited. And this is this gap that we are covering through our work, through looking into potentials but also risks of applying AI in peacebuilding. We are looking at peace, not just as absence of violence and violent conflict, but as thriving society, society where everyone can thrive. And we turn to using the definition of peacebuilding as a strategic activity to solidify peace and avoid violent conflict, to provide the tools for building something more than just the absence of war, and then to strengthen capacities for a conflict management tool and to lay the foundations for sustainable peace and development.

8:56 We aim at working with peace builders, and we want peace builders to be more informed about the potentials of AI, but also more informed about risks. We want them to be ready to build their capacities to use these tools, but also to speak up and act when needed. But we are also talking with AI and data scientists, so we want them to become a little bit of peace builders themselves and to commit to use of AI for creating lasting peace. So looking into the latest developments in technologies in big data, we see that they're bringing

Transcribed by https://otter.ai

transformational changes to societies including some changes to peace, justice and human rights. But at the same time, they're bringing a vast number of complex ethical questions and challenges. So greater availability of data, increased adoption of Big Data practices, utilization of machine learning, deep learning of so-called blackbox AI. They have all brought more attention to ethical challenges, especially around collection, use, processing of big data, around data ownership, around sovereignty issues, around consent of those whose data is being processed, about trustworthiness of AI systems, about risks of reinforcing existing gender, racial and other biases, or disregarding human rights and privacy, accountability, security, transparency, safety, and many other issues.

10:50 What our work with AI for Peace Community of Practice showed is there are more and more organizations from the peacebuilding field, even those who have limited resources, and often because they have limited resources, and although often lagging behind other sectors in technical and resource capacities, they're already looking into the power of algorithms, and how machine learning can potentially advance their work and help populations in need. So while being at this stage, very early stage of development and as the data science and AI become increasingly accessible and used across the peacebuilding sector and in sustaining peace work, we are asking this question of how can we ensure ethical work is being done?

11:51 So how do we approach ethics in AI for peace? So ethics is not new to peacebuilders, however the majority of peacebuilding ethics were concentrated around normative ethics responding to questions such as who has the agency or who should have agency in peacebuilding work? What ends should peacebuilding pursue? And what means will ensure that peace building is done right? Should the international community engage in peacebuilding? Or should it be exclusively a local activity? Who should be involved? What constitutes good ends for peacebuilding? And how can peace building be done right? What we are specifically looking into is the field of AI ethics in peacebuilding and all the added challenges brought by AI field, and by the so-called Algo Age, right, the age of algorithms. We also look how we can build on the lessons of ethics peacebuilding work, but also fill in the gaps created with specifics of the algorithms and algorithmic age.

13:16 So, how do we define AI ethics is all not that different than we're taking here the definition of Alan Turing Institute that and this is a quote saying that: “Ethics is a set of values, principles, techniques that employ widely accepted standards of right and wrong to guide moral conduct in the development and use of AI technologies”. And we are adding “in sustaining peace” or according peacebuilding work. So what is our community of practice saying when we started doing interviews with different organizations, who are already applying data science in peacebuilding work and asking how they're actually embedding ethics in their own work. So those are some of the quotes from the interviews that “ethics is more of a process rather than a destination” that organizations are aiming at. It is “a place to pause and ask questions” about projects they're doing. And it is also an “opportunity to think about unintended consequences”.

14:29 So what we did is actually first looking into the existing principles and guidelines, and seeing how do they actually fit into the work and goals of the peacebuilding work. Can we actually build on the work of the existing AI ethics field? And in recent years with raising awareness about the importance of AI responsible practices, a growing number of organizations started adopting those AI principles and frameworks. So our Algorithmic Watch initiated an AI ethics guidelines inventory in which more than 160 AI ethics guidelines have been collected. More Algorithm Watch also reported about is that majority of these guidelines are mostly containing general promises but there are no concrete recommendations of how to actually bring these principles to practice, which was also an

Transcribed by https://otter.ai

insight that our interviews provided from majority of organizations we talked with that they really struggle in building in bringing different principles to practice.

15:58 Then similar work has been done by Partnership on AI, which is currently in the process of conducting validation of different AI explainability tools with developers and users of those tools, determining how useful they actually find them in practice. And they also reviewed and are currently preparing to publish a database of around 150 tools. They're not the only one. There is an Organisation for Economic Cooperation and Development (OECD) also focusing on creating an interactive database for various AI tools, practices and approaches for implementing trustworthy AI, so-called tools for trustworthy AI to help practitioners to determine which tools they can actually use in their work.

16:54 So talking with peacebuilding practitioners, we realize that first of all, some of them do lack awareness about the existence of these tools. On the other side, those who are aware about the existence are struggling with the so-called Choice Overload. So we realize there are too many options that lead people to make other poor choices with undesirable outcomes or, or even make them completely avoid making choices. Another thing coming from this research of Algorithm Watch is that despite the inflation of guidelines for the ethical conduct of AI research, there is no shortage of reports of unethical applications of AI. So our research also showed another problem that peacebuilding organizations are facing with and it's sort of paradoxical that we this choice overload and such a big amount of AI ethics available and tools, peacebuilders are pointing out that there is still a gap in research and tools covering specific ethics of artificial intelligence in peace. So they struggle to see how these tools will actually be practically useful in their own work. And we saw something similar in terms of what is publicly available out there. For example, the *Oxford Handbook of Ethics of AI* offers a great overview of different cases and applications of AI, including the ethics of AI in transport, in biomedical research, in public health, in law, in criminal law, in migration management. But there's none of these resources refers to the specific field of peacebuilding.

19:10 When we look at what is happening in the field, we see that there is a dire need of tools like this that could ensure more ethical work in conflict and fragile countries. This is just a reminder of several issues that were widely covered in news, in media, in research such as Facebook being misused to incite violence and ethnic cleansing in Myanmar. Even the report, the internal report commissioned by Facebook that found that the company failed to keep its platform from being used to foment division and incite offline violence in Myanmar in a very fragile situation. This is not the only unintended consequence or risk brought by algorithms and data. This is another example of refugees being exposed to further risks. This is an example of UNHCR collecting and sharing personal information from again ethnic Rohingya refugees with Bangladesh, which then shared this data with Myanmar to verify people for possible repatriation. So potentially putting people in additional risk.

20:45 Afghanistan is the latest reminder. So this is just a year ago of this urgent need to understand all aspects of emerging new technologies with humanitarian peace building and development work. As the Taliban seized control of major cities a year ago, people were concerned that the group could use data, social media and other online information to identify citizens who previously worked for the Afghan security forces, for civilian government, or for foreign organizations. So people really struggle to understand the full scope of the situation left behind by international actors and their data driven projects. Activists were stressing the risks to Afghans

Transcribed by https://otter.ai

created by all actors who were previously providing assistance through data and digital tools, and they were calling for greater protection.

21:52 So this is just another reminder about the need to establish a mechanism that ensures responsibility of actors in conflict and fragile settings to protect vulnerable populations from both intended and unintended consequences. And this is the example that is showing us that risks go beyond digital security concerns and include serious physical and security risks, which is again, another reason to really make sure that we understand how to ensure ethical practices in fragile countries that do not harm populations that are already in risk.

22:42 So there are many organizations which is good started recognizing the need to develop these ethical and responsible approaches. So what we are looking at in our work is also existing not only AI ethics tools, but specifically the tools in the humanitarian sectors such as the OCHA Data Responsibility Guidelines, which would be the closest to our field, to the peacebuilding field. And the Center's data responsibility work is focused on developing guidance processes and practices for how UN OCHA specifically handles data as the coordinator of the entire humanitarian response. And one of the direct results of these efforts is the publishing of the OCHA Data Responsibility Guidelines in 2021, establishing those principles, processes and tools to support or sustain OCHA's data work.

23:52 Another important tool we are bringing to our field is ICRC's data protection framework, providing individuals with the protections in according with international standards. Yet again, another tool from the humanitarian sector comes from the Humanitarian Data Science and Ethics Group. And many of these ethics frameworks that were published in 2020 have brought the ethical concerns closer to our field as well. Especially as most of the applications are happening in crisis situations right, in conflict affected countries, and they're often affecting, as I mentioned, already vulnerable populations. So this uniquely, sensitive nature of peacebuilding brings heightened ethical considerations and challenges. This work to advance the field of ethical technologies, and lately, ethical AI has paved the way for practices and principles also in peacebuilding and prevention fields. I mentioned a couple of them, but there are numerous humanitarian agencies that have published their own data protection manuals. But it's really important to also understand that the ethical challenges of exponential technologies go beyond data protection and security measures, and should include additional focus on issues specific to, to peacebuilding.

25:46 So this is why the recommendation we're bringing to the table and what we are working on is building on the existing good practices from the humanitarian and broader AI ethics field, and then upgrading those with the specifics of peace builders, of the needs that peace builders have in the field to actually provide ethical guidelines for AI for peace. And failing to operationalize AI ethics, for example in business, can lead to loss of resources, inefficiency in product development and deployment, and in many cases, reputational risks. But I want to emphasize again, for peacebuilding field, it's much more than the reputational risk, it's actually the impact on the people, on the populations that live in those countries, in fragile and conflict areas that are in question here.

26:53 What our research showed is also there, as I mentioned, many ethical tools out there can be overwhelming for organizations that are often under-resourced in their work. We are also recognizing the context in which we are working as peacebuilders is quite different. And also the basic knowledge of AI ethics concepts and tools was very diverse among different stakeholders we interviewed, while minority were experts and even creators of some

Transcribed by https://otter.ai

of these tools, majority was actually lacking the basic literacy of what already exists in AI ethics. So, we identified this gap. And we are offering this ethical guidelines for AI for peace as sort of a centralized resource that can explain some of these concepts, but also summarize some of the tools available for organizations conducting AI driven projects.

28:11 So this would be a living knowledge, a living knowledge base recognizes that things are also changing, technologies are changing, and new technologies new AI approaches are being developed. So we need to keep our awareness and knowledge about them. And it's also a resource for anybody working in the AI for peace field, or even organizations who are only exploring how they can enter this field to potentially advance their mission. This resource would also share best practices for approaching data science and AI projects in the peacebuilding field with an ethics lens.

29:03 There is another question in the dimension of this work that we are bringing to the AI ethics field, and this is where AI for peace community can actually help. So what we are recognizing is that there are certain approaches that peacebuilders were using in their peacebuilding, in their ethics for peacebuilding, work that can be potentially upgraded to be useful for the age of algorithms or for any organizations that are utilizing algorithms in their work. And those are conflict sensitivity and do no harm approaches. So we also, so we want to recognize that peacebuilders are not only catching up with this rich and extensive debate on AI principles, on transparency, accountability, privacy, ethical and human rights approaches of algorithmic decision making, but they also have something to contribute, especially their experiences in sensitive conflict and violence affected areas and other high risk settings that can be beneficial for how to design, develop and implement data driven approaches.

30:26 Our work aims at looking into the specific benefits that can be achieved through the conflict sensitivity, and do no harm approaches and their obligations in the Algo Age. So what is conflict sensitivity? It's actually a conceptual framework, and a set of tools that can help stakeholders navigate complexities of doing projects in conflict affected areas. So conflict sensitivity does not have to be applied only by peacebuilders, right? It can be utilized by any stakeholder as guidelines for tackling complex complex situations, and diminishing the potential negative impacts of their products on violence and conflict. So, conflict sensitivity allows actors to understand first the context in which they operate, to understand the interactions between their activities and the context, to take some specific steps to minimize the negative impacts of their operations, and to take steps also to maximize the positive effects and impacts of their operations. So to contribute to peace, not only to make sure not to do harm, but to actually sustain peace and improve peace. So in complex, in complex settings, good intentions can sometimes make a situation worse and seen through this lens of conflict sensitivity is based on the premise that any activity, any project is not always neutral, especially when applied to areas affected by violent conflict. Often the product, the outcome of a project, of an activity, can be misused to strengthen one of the sides of the conflict or weaken the other side. So the possibility of harm actually increases when we bring the project to the conflict area and when we fail to identify and address the effects of otherwise well intended actions.

32:58 So with this, we want to conclude that AI technology is here to stay, and while some approaches will show less effective for achieving peacebuilding goals, other AI tools will appear and grow in the future. As we saw with the development of natural language processing, neural networks, deep learning that brought some of the new potential applications in peacebuilding field. The success or failure of sustaining peace with these tools, the possibility of stopping unintended consequences and ensuring accountability depends on whether our field is

Transcribed by https://otter.ai

committed, and investing in embedding ethics in our work by design as a natural first step in developing, designing and deploying these tools, rather than some, after thought to be dealt with after the deployment of the technology. So we hope that the ethical guidelines for AI for peace will contribute to this goal, and we invite you if you're interested to learn more about the guidelines or about the other products, outcomes of this project, you visit our website: aiforpeace. And you can also follow us on Twitter at AI4Peace and join our AI for Peace LinkedIn group. Thank You.

Transcribed by https://otter.ai