United Nations University
Centre for Policy Research

Predictive Technologies in Conflict Prevention: Practical and Policy Considerations for the Multilateral System

Discussion Paper, June 2023

Eduardo Albrecht

UNU CPR

About the author
Eduardo Albrecht is a Senior Fellow (Non-Resident) at United Nations University Centre for Policy Research and Associate Professor in the Department of Social Sciences at Mercy College in New York.

Acknowledgements
The author would like to thank Dr Adam Day, Head of the Geneva Office of the United Nations University Centre for Policy Research, for thoughtful review of earlier drafts of this paper. The paper was compiled with research assistance from Apna Balgobin, Davina Resto, and Aurora Rudd.

Disclaimer
The views and opinions expressed in this paper do not necessarily reflect the official policy or position of the United Nations University.

UNU-CPR Discussion Paper Series
UNU-CPR Discussion Papers provide an in-depth analysis of current and emerging issues of strategic importance to the UN and Member States. They reflect UNU-CPR research priorities, identify potential solutions to key challenges, and are designed to generate discussion and comment.

ISBN: 978-92-808-6590-5

Citation: Eduardo Albrecht, "Predictive Technologies in Conflict Prevention: Practical and Policy Considerations for the Multilateral System," *UNU-CPR Discussion Paper* (New York: United Nations University, 2023).

Predictive Technologies in Conflict Prevention: Practical and Policy Considerations for the Multilateral System

Discussion Paper, June 2023

Eduardo Albrecht

Contents

1. Introduction

Conflict prevention has become increasingly central to the UN's approach to insecurity and instability, and this shift has brought a greater reliance on data capture technologies to identify and analyse recurrent conflict patterns and forecast potential crises.

However, the use of these technologies also poses several challenges. For instance, some possess inconsistencies in the quality of data across space and time which undermine the ability of the UN to accurately predict conflict trends. There are further concerns that data capture technologies used for predictive purposes could have serious security ramifications and could run afoul of mainstream data privacy standards.

Additionally, there is always the risk that technologies could be adapted for surveillance purposes in violation of human rights, and there are further concerns regarding 'automation bias' – a human tendency to be less critical of suggestions made by automated decision-making systems which could result in an over-reliance on predictive technologies, complicating an organization's ability to respond effectively in fast-moving and emergent conflict scenarios.

Although fraught with various growing pains, these technologies are nonetheless destined to become more central to the UN's prevention toolkit. This discussion paper engages various aspects of predictive technologies in conflict prevention and begins to outline an approach to their adoption that mitigates some of the challenges related to their uptake and application.

The first section reviews steps UN bodies have taken to use predictive technologies for conflict prevention. Next, it considers Member States' role in advancing these technologies in the multilateral system, a field known as 'peacetech.' The last section looks at ethical aspects, including a review of the possible unintended consequences of deploying such technologies. Taken together, these sections explore what Member States could do to best position the multilateral system for the effective and ethical uptake of predictive technologies for conflict prevention and management.

The central recommendation is to adopt the precautionary principle when considering deployment of predictive systems. Member States should work alongside the policy research community and peace operators in the field to spearhead a common process for this precautionary approach. For instance, models could be trialled in a safe and controlled environment – for example the UN Futures Lab, or the UN Departments of Political and Peacebuilding Affairs (DPPA) and Peace Operations (DPO) – before attaining 'approval' and guidance for wider use. This would give those involved in conflict prevention and management work in the UN System the confidence to use the technologies in the field without having to worry about causing any unintentional harm.

2. Predictive Technologies in the UN Conflict Prevention/ Response Architecture

The UN's approach to insecurity and instability increasingly emphasizes conflict prevention. This is most evidenced in the 2016 Sustaining Peace resolutions and the 2018 reform of the UN peace and security architecture, driven by the Secretary-General's call to make the entire system "work for prevention".[1] Part of this push for prevention has meant a greater use of "data capture technologies and intelligence collection to map and understand recurrent conflict patterns and forecast potential crises".[2] Since 2014, when the Expert Panel on Technology and Innovation in UN Peacekeeping observed that "information is a political resource", myriad open-source and proprietary data streams have been tapped into for conflict prevention efforts. As the panel noted: "[V]oice, video, and data from commercial satellites, sensor networks, and other technical feeds are available and need to be used by UN decision-makers".[3]

The technologies employed to crunch these data are typically internally generated or sourced via private companies. Various UN entities have built information repositories and analytic capacities to support prevention mandates. Other UN entities have partnered with external private actors to supply predictive analytics based on social media, earth observation, and many other types of digital information generated from routine human behaviour. Often, the UN uses a mix of both in-house and external tools. Whatever the source, when appropriately brought together, these technologies can give advance insight into future political and civil unrest. Given that to predict is to possibly prevent, the technologies are becoming ever more relevant to the UN's ability to fulfil its peace and security mandates.

Steps have been taken within the UN system to accelerate the use of these technologies. Forward-looking 'pockets' within the UN, like the DPPA's Innovation Cell, have been exploring multiple scenarios for AI and "its potential to be used for the noble purposes of peace and security", which "could revolutionise the way of how we prevent and solve conflicts globally".[4] Similarly, the DPPA's Mediation Support Unit (MSU) has developed a Digital Technologies and Mediation Toolkit that explains how data analytics and machine learning could "be used for the purpose of conflict analysis, early warning, [and] prediction of conflict". The toolkit further clarifies that mediators could "generate predictions about what conflict stakeholders will do, when and where".[5] The UN Global Pulse, too, has been advancing work around using data for foresight "to become more data-informed and anticipatory in not only strategy but also planning and implementation".[6]

A leading in-house example is the United Nations Development Programme's (UNDP) Crisis Risk Dashboard, or CRD. Started in 2016 under the direction of the Crisis Risks and Early Warning team, the CRD "ensures that relevant and updated data is readily available to support processes of monitoring, analysis, and formulation of anticipatory measures and responses". The team is leveraging innovative technologies as they become available. For example, at a time of increasingly hostile online discourse, several dashboards "are being developed to use sentiment analysis, machine learning, and predictive technology to find patterns and detect early warning signals" as the basis for more effective programmes to mitigate violence.[7]

---

1 Eleonore Pauwels, "Artificial Intelligence and Data Capture Technologies in Violence and Conflict Prevention: Opportunities and Challenges for the International Community," Policy Brief (New York: Global Center on Cooperative Security, 2020).

2 Ibid.

3 Martin Wählisch, "Big Data, New Technologies, and Sustainable Peace: Challenges and Opportunities for the UN," *Journal of Peacebuilding & Development*, Vol. 15 Issue 1 (2020): 122–123.

4 Daanish Masood and Martin Wählisch, "Future Wars Will Be Waged With Robots. But So Might Future Peace," *EuroNews*, 28 March 2019.

5 The UN Digital Toolkit is available at https://peacemaker.un.org/digitaltoolkit. To the toolkit creators' credit, they emphasize that these technologies must be handled with care, as output "can be shaped by the cognitive and social biases underlying the programming algorithms". These biases may, in turn "engender discrimination towards traditionally excluded groups and vulnerable communities". They correctly point out that "context and technical experts would be required to correct and adjust the machine learning process and contribute their knowledge and analysis to improve accuracy".

6 Amy Lynn Smith, "Using Imagination, Information, and Insight to Prepare the UN for the Future," *UN Global Pulse*, 22 February 2022, https://www.unglobalpulse.org/2022/02/using-imagination-information-and-insight-to-prepare-the-un-for-the-future.

7 Corrado Scognamillo and Jonas Gutschke, "Understanding the World through Data: United Nations Development Programme," *UNDP Blog*, 9 November 2021, https://www.undp.org/blog/understanding-world-through-data.

Some internally generated information repositories fall short of becoming predictive tools. An example of this is the Situational Awareness Geospatial Enterprise, or SAGE. Experts have described how the tool "allows UN military, police and civilians in UN peace operations (both UN peacekeeping operations and special political missions) to log incidents, events, and activities". They also note that this treasure trove of conflict data, combined with machine learning, could potentially "offer a giant step forward in the predictive capability of the UN, and hopefully be translated to preventive action on the ground", and detail how predictive analyses could, in principle, generate "risk maps" where a "colour coding of administrative districts indicates the probability of events of interest like armed clashes between the main warring parties, communal violence, or violence against civilians".[8]

Yet, to date, this has not occurred. There are a number of headwinds to tools like SAGE being used in this way. A survey of a wide array of digital technologies used in peace operations found that SAGE possessed inconsistencies in the quality of the data across space and time. Additionally, the database only records information about events (as opposed to individuals or contextual factors). These characteristics may undermine the ability of the UN to accurately predict trends.[9] Moreover, serious potential privacy concerns could arise if UN databases were used for predictive purposes.[10] UN analysts often rely on local informants to gather data in conflict-affected areas, as this information is not always available from other sources. Parties to a conflict may attack individuals thought to have given information to the UN. This way, "civilians who are already at risk can face new threats if their personal information is disclosed or reidentified".[11] This would particularly be the case if combatants knew that such information was used to inform preventive approaches that may result in an unfavourable distribution of UN peacekeeping resources. In sum, turning a database into a predictive resource has serious security ramifications and could run afoul of mainstream data privacy standards.

Similar concerns have been expressed towards a tool developed by the UN Global Pulse, called Qatalog, and another developed by the DPPA, called Sparrow. Qatalog uses AI-based language processing technology to automatically 'listen' to public radio talk shows and 'read' public Twitter streams in 39 different languages across the globe. The tool allows UN analysts to personalize machine learning text classifications and extract useful information on trends.[12] Sparrow is a social media scanning tool that separates noise from authentic conversations. We know that a portion of online conversations is intended to sway public opinion in favour of this or that political agenda – often using incorrect facts, sponsored by foreign actors, and fuelled by automated systems or 'bots.' Sparrow addresses this concern as it "allows UN desk officers in New York and in field missions to rapidly analyse Twitter data and separate 'noise' created by bots on social media from authentic political speech", thereby capturing actual sentiment trends in the population.[13] These harvested data, alongside other information on conflict events, could potentially be mined for predictive signals. However, as it stands, Qatalog and Sparrow do not formulate predictions but rather monitor current and past trends.

UN-contracted private companies supplying data analytics tend to be less lead-footed regarding their products' predictive capacity. Observers note the use of third-party social media analysis tools such as Dataminr and Predata. These platforms, "in use in pockets of DPO/DPPA Headquarters and missions", claim to function as conflict early warning tools. Yet, several factors limit their actual utility in conflict prevention. For example, Internet coverage is likely limited in areas affected by conflict, which can create data blind spots. Also, disparities in the local usage of social media can make meaningful analyses across populations difficult. As such, "few, if any, social media analysis tools can offer predictive insights with sufficient geographic precision in peacekeeping contexts to be tactically actionable".[14] A similar doubt has been cast on the capacity of automated data extraction algorithms, such as

8 Allard Duursma and John Karlsrud, *Predictive Peacekeeping: Opportunities and Challenges* (Oslo: Norwegian Institute of International Affairs, 2018).

9 Dirk Druet, *Thematic Research Paper for the DPO Peacekeeping Technology Strategy: Enhancing the Use of Digital Technology for Integrated Situational Awareness and Peacekeeping-Intelligence* (New York: United Nations Department of Peace Operations, 2021).

10 Allard Duursma and John Karlsrud, *Predictive Peacekeeping: Opportunities and Challenges* (Oslo: Norwegian Institute of International Affairs, 2018).

11 Ibid. The UN Peacebuilding Support Office, in fact, explicitly recommends that "any intervention using these technologies must be mindful of the operational and ethical risks associated with using data that can be linked to personally identifiable information". See Lorena Escobal, Kelsey Finnegan, Hayung Kim, Hyunwoo Park, James Schalkwyk, and Asher Zlotnik, *Big Data for Peace and Security* (New York: Columbia SIPA and UN Peacebuilding Support Office, 2018), pp. 5–6.

12 "Qatalog," UN Global Pulse, last accessed on 30 January 2023, https://www.unglobalpulse.org/microsite/qatalog/.

13 Politically Speaking, "'We Use Technology, Not the Other Way Around' - Social Media and Political Analysis," Medium, 30 September 2021, https://dppa.medium.com/we-use-technology-not-the-other-way-around-social-media-and-political-analysis-e97706ba0465.

14 Dirk Druet, *Thematic Research Paper for the DPO Peacekeeping Technology Strategy: Enhancing the Use of Digital Technology for Integrated Situational Awareness and Peacekeeping-Intelligence* (New York: United Nations Department of Peace Operations, 2021), p. 10.

web scraping and signal detection based on social media, to "forecast low-probability conflict events with high temporal and spatial accuracy".[15]

We must add that even when and where the tools do work, uptake by peace operators is not always a given. Kristian Hoelscher and Jason Miklian, senior researchers at the Peace Research Institute Oslo, talk about an 'expertise gap' that occurs when private technology start-ups jump into the global peacebuilding space. Recent years have seen a plethora of such start-ups surface, "often through government and philanthropic funders who believe that cutting-edge technologies can help mitigate political evils". The trouble, they point out, is that "it leads to an expertise gap as start-ups launch peace tools without employing existing peacebuilding knowledge – or worse, don't think that such knowledge is needed at all … then, experts dismiss the well-meaning initiatives as being hopelessly naive to complex conflict realities".[16]

One member of the Innovation Cell at DPPA noted that "new technologies cannot be a panacea for any analytical question in conflict prevention or any operational challenge in peacemaking". According to the expert, the most advanced technologies can only leverage "diplomatic efforts to a certain extent: Personal experience and gut feeling for political nuances cannot be replaced by machines, yet".[17]

In addition, there is always the risk that "technologies built for lawful use can easily be adapted to facilitate surveillance in violation of human rights principles".[18] An excess of caution may rightly dampen the zest to pick up certain tools even when they are arguably within reach. Additionally, these technologies create a risk of 'automation bias,' where humans have a documented tendency to be less critical of suggestions made by automated decision-making systems.[19] This bias can result in an over-reliance on predictive technologies (and an over-confidence in their output), which may complicate an organization's ability to respond effectively in fast-moving and emergent conflict scenarios. These and related concerns (see Figure 1 for an overview of these obstacles) have likely contributed to curbing efforts to turn tools like SAGE, Qatalog, and Sparrow from purely monitoring ones to predictive ones too and generated appropriate scepticism toward tech companies' out-of-the-box software solutions.

Figure 1.

15 Allard Duursma and John Karlsrud, *Predictive Peacekeeping: Opportunities and Challenges* (Oslo: Norwegian Institute of International Affairs, 2018), pp. 6–8.

16 Kristian Hoelscher and Jason Miklian, "Can Innovators be Peacebuilders? A Peace Innovation Action Plan," *Global Policy*, 1 August 2017, https://www.globalpolicyjournal.com/blog/01/08/2017/can-innovators-be-peacebuilders-peace-innovation-action-plan.

17 Martin Wählisch, "Big Data, New Technologies, and Sustainable Peace: Challenges and Opportunities for the UN," *Journal of Peacebuilding & Development*, Vol. 15 Issue 1 (2020): 122-126, p. 123.

18 Eleonore Pauwels, "Artificial Intelligence and Data Capture Technologies in Violence and Conflict Prevention: Opportunities and Challenges for the International Community," Policy Brief (New York: Global Center on Cooperative Security, 2020), p. 11.

19 Ibid., p. 16.

While fraught with various growing pains, the technologies described so far are nonetheless destined to become more central to the UN's prevention toolkit. Over the past two decades, we have seen significant advances in the capacity to forecast conflicts related to government instability, climate change, terrorism, political protest, and war.[20] Many formidable tools have been developed, and many more are coming. The implications for international development and security are wide-ranging.[21] The next section will review these tools and the role of Member States in advancing their utility to conflict prevention and management.

20 For a review of these applications, see Robert Trappl (ed), *Programming for Peace: Computer-Aided Methods for International Conflict Resolution and Prevention*, (Springer: 2006); Jack Goldstone et al. "A Global Model for Forecasting Political Instability, *American Journal of Political Science*, Vol. 51 Issue 1 (2010): 190-208; Sean O'Brien, "Crisis, Early Warning and Decision Support: Contemporary Approaches and Thoughts on Future Research," *International Studies Review* Vol. 12 Issue 1 (2010): 87-104; Håvard Hegre, Joakim Karlsen, Håvard Mokleiv Nygård, Håvard Strand and Henrik Urdal, "Predicting Armed Conflict, 2010–2050," *International Studies Quarterly* Vol. 57 Issue 2 (2013): 250-270; Franceso Mancini, ed. *New Technology and the Prevention of Violence and Conflict* (New York: International Peace Institute, 2013); Weisi Guo, Kristian Gleditsch, and Alan Wilson "Retool AI to Forecast and Limit Wars," *Nature Comment*, 15 October, 2018, https://www.nature.com/articles/d41586-018-07026-4; Ravi Gupta and Hugh Brooks, *Using Social Media for Global Security* (Wiley, 2013); Christian Reuter, ed. *Information Technology for Peace and Security: IT Applications and Infrastructures in Conflicts, Crises, War, and Peace* (Springer, 2019); and Andre Petheram, Eleanor Shearer, Richard Stirling, and Tom Westgarth, *Fakes, Files, and Facial Recognition* (Oxford: Oxford Insights, 2020).

21 Siddhartha Raja, Tatiana Nadyseva, Roku Fukui, Rachel Firestone, and Michael Minges, "People and Data," in *Information and Communications for Development: Data-Driven Development* ed. Boutheina Guermazi (Washington, DC: World Bank Group, 2018).

3. The Role of Member States in Advancing Predictive Capacity for Peace Operations

In 2017, the Swiss Foreign Department of Federal Affairs and the Canton of Basel-Stadt launched the Basel Peace Forum. The forum intends to "inspire new and unconventional ideas for peacebuilding". To this end, decision makers, diplomats, academics, and civil society leaders meet yearly to "rethink peace". As early as 2017, linkages between peace, artificial intelligence, and risk analysis took centre stage. David Lanz, Head of the Mediation Program at Swisspeace, wrote in a summary of the workshop exploring these linkages that "research around AI needs to be refocused to how the technology can promote peace, rather than wage war".[22] This is no small point. Typically, the vast majority of State investment in AI and related technologies has gone to advance nations' defence and intelligence apparatuses. Nonetheless, and to their credit, a number of States have steadily increased investment in AI-fueled early warning projects that can potentially help prevent armed conflict, politically- and ethnically-motivated violence, and mass atrocities in various geographies.

Much of this investment is carried out in partnership with academia.[23] For example, the Violence Early Warning System (ViEWS) was created by researchers at Uppsala University and the Peace Research Institute Oslo. It can automatically identify future instances of violence in Africa. ViEWS is publicly available, data-driven, and "generates monthly probabilistic assessments of the likelihood that fatal political violence will occur". Its predictions can see up to three years into the future for each 55x55 km grid cell throughout the African continent. The work is funded by the European Research Council, the Swedish Research Council, the UN High Commissioner for Refugees, and the UK Foreign and Commonwealth Office, among others.[24] A similar example is Conflict Forecast. Investigators here have developed a model that can predict outbreaks of internal armed conflict by automatedly parsing through "millions of newspaper articles" since 1989. The project predicts conflict up to a year in advance in 180 countries worldwide. It is funded by the UK Foreign and Commonwealth Office and the Spanish National Research Council.[25]

Another case of successful academic-government cooperation is the Armed Conflict Location and Event Data project (ACLED). Its prediction tool, Conflict Pulse, predicts whether there will be an increase in the number of conflict events for a given actor compared to the previous week. The forecasts are made using ACLED's event data – "information on the dates, actors, locations, fatalities, and types of all reported political violence and protest events around the world" – which is a handy resource for practitioners and scholars.[26] The project received funding from the University of Texas at Austin, the European Research Council, the US Department of State, the German Federal Foreign Office, and the Dutch Ministry of Foreign Affairs. Multilateral support has come via the International Organization for Migration (IOM) and the Complex Risk Analytics Fund (CRAF'd).[27]

The UN-hosted CRAF'd, launched in collaboration with the World Bank in 2021 and supported by the Governments of Germany, The Netherlands, US, and Finland, is poised to become a key player in the field and allow multilateral organizations to play a more significant role in advancing this kind of work. The fund is a multilateral financing instrument that seeks to "expand shared capabilities for using data to better anticipate, prevent, and respond to

22 David Lanz, *New Technologies to Prevent Conflict and Build Peace: Critical Reflections of the Basel Peace Forum Workshops on AI, Warfare, Ethics* (Basel: Basel Peace Forum, 2017).

23 Cooperation with academia is quite common in the field. Also, the UN DPPA's Innovation Cell, for example, cooperates with researchers at Stanford University to explore the correlation between depleting groundwater and civil unrest and MIT to better sharpen AI tools that monitor social media conversations occurring in dialects and languages around the world. See Dalvin Brown, "The United Nations Is Turning to Artificial Intelligence in Search for Peace in War Zones," *The Washington Post*, 23 April 2021; and Politically Speaking, "Getting to Grips with New Tech in Prevention and Peacemaking," *Medium*, 22 November 2019.

24 "About VIEWS," VIEWS, last accessed 30 January 2023, https://viewsforecasting.org/about/; and "ViEWS," Uppsala Universitet, last accessed 30 January 2023, https://www.pcr.uu.se/research/views/; See also Tate Ryan-Mosley, "We Are Finally Getting Better at Predicting Organised Conflict," *MIT Technology Review*, 24 October 2019.

25 "The Project," Conflict Forecast, last accessed on 30 January 2023, https://conflictforecast.org/about.

26 "Early Warning Research Hub," ACLED, last accessed on 30 January 2023, https://acleddata.com/conflict-pulse/; and "About ACLED," ACLED, last accessed on 30 January 2023, https://acleddata.com/about-acled/.

27 Ibid.

complex risks in fragile and crisis-affected settings"[28] and "to spur anticipatory action before disasters unfold".[29] To this end, it invests in "analytics, including predictive models, as well as methods for the analysis of social media, geospatial, [and] other data for crisis anticipation, prevention, and response".[30]

Other initiatives are funded by defence agencies. The Turing Group, in partnership with UK Government defence and security agencies, developed a technology called the Global Urban Analytics for Resilient Defence, or GUARD. GUARD enables peacekeepers to predict where urban conflict will likely break out 12 months in advance. Veronica Wardman, a technical partner for the project, states that "different populations and cultures have different dynamics, trigger points, and strategic influence, which must be appreciated and taken into account". One of the main innovations of GUARD is that "it seeks to unpick and understand this space using the latest cutting-edge tools and techniques".[31] Similarly, the US military's Integrated Crisis Early Warning System (ICEWS) can automatically monitor and forecast events that could affect national security interests.[32]

In other cases, projects may start with public research funds but then find further support from the private sector. The free, open source resource called the Global Database of Events, Language, and Tone, or GDELT, is a good example. Here, work funded in part by US National Science Foundation grants contributed to creating a news monitoring and forecasting capacity, which was then expanded via support from Jigsaw, a technology incubator created by Google. GDELT "monitors the world's news media from nearly every corner of every country in print, broadcast, and web formats, in over 100 languages, every moment of every day". All global news is translated in real-time into English and categorised according to hundreds of event types, thousands of emotions, and millions of themes. GDELT's

Leading Conflict Prediction Initiatives

States have increased investment in AI-based early warning projects, often in partnership with academia and in collaboration with other States.

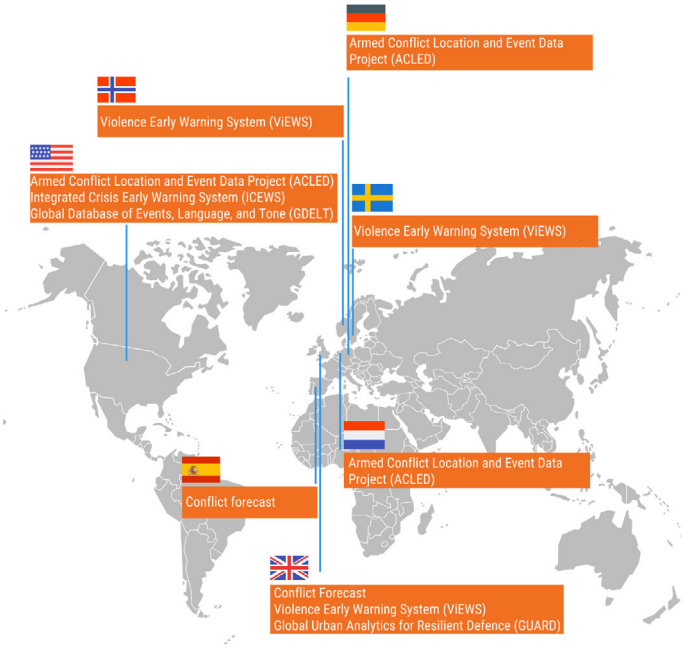

Figure 2.

Risk Assessment and Global Trends solution offers visibility into emerging conflict trends by summarizing major emerging risk trends in the last 48 hour news cycle and comparing them to the previous 48 hours.[33]

These are only some of the many projects in what has sometimes been referred to as the 'peacetech' field. For a summary of where leading conflict prediction initiatives are located see Figure 2. The term, in vogue with the establishment of representative non-profits like Build Up, Peacetech Lab, and JustPeace Labs,[34] collectively refers to

28 "What is CRAF'd?" CRAF'd, last accessed on 30 January 2023, https://crafd.io/.

29 United Nations, "*United Nations and Partners Launch Complex Risk Analytic Fund to Unlock Power of Data for Crisis Action*," 13 October 2021.

30 "CRAF'd will join up investments to unlock scale and potential," CRAF'd, last accessed on 30 January 2023, https://crafd.io/what-we-support.

31 "Predicting Conflict – a Year in Advance," The Alan Turing Institute, last accessed 30 January 2023, https://www.turing.ac.uk/research/impact-stories/predicting-conflict-year-advance; See also "Artificial Intelligence (Safe and Ethical)," The Alan Turing Institute, last accessed 30 January 2023, https://www.turing.ac.uk/research/research-programmes/artificial-intelligence-ai/safe-and-ethical.

32 Sean O'Brien, "Crisis Early Warning and Decision Support: Contemporary Approaches and Thoughts on Future Research." *International Studies Review*, Vol. 12 Issue 1 (2010): 87–104.

33 "The GDELT Story," The GDELT Project, last accessed on 30 January 2023, https://www.gdeltproject.org/. For an example of academia-civil society collaboration see the Early Warning Project (EWP). This system can assess the likelihood of mass atrocities in countries worldwide up to two years in advance. EWP is a joint initiative of the Simon-Skjodt Center for the Prevention of Genocide at the United States Holocaust Memorial Museum and the Dickey Center for International Understanding at Dartmouth College, https://earlywarningproject.ushmm.org/. For a system with a similar objective but with forecasting capacity still under development, see the Sentinel Project's Early Warning System, https://thesentinelproject.org/what-we-do/early-warning-system/.

34 See "We transform conflict in the digital age," Build Up, last accessed on 30 January 2023, https://howtobuildup.org/; "Putting the right tools in the right hands," PeaceTechLab, last accessed on 30 January 2023, https://www.peacetechlab.org/; and "Advancing peace and human rights through the power of technology," JustPeace Labs, last accessed on 30 January 2023, https://justpeacelabs.org/.

those initiatives operating at the intersection of data, human rights, and peacebuilding. For more information on this space, readers may want to peruse a New York University Center on International Cooperation (CIC) resource called the *Ecosystem Map: Data for Peacebuilding and Prevention*. The Ecosystem Map is an interactive digital tool that surveys all existing global organizations (ranging from civil society, to government, to the private sector) working with data for peacebuilding.[35] Developments in the field have been so fast and varied that it can be hard to keep track of where innovation is occuring, making this tool especially valuable. For example, data shows that civil society, mostly in the Global North, is leading the charge in peacetech (see Figures 3 and 4). Paige Arthur, a fellow at the center, and Branka Panic, a visiting scholar, make a strong case for the Ecosystem Map in a CIC publication that describes how a centralized repository of different data technologies is essential to peacebuilding in the digital era.[36]

Somewhat less established is State funding of research around the policy and governance implications that naturally surface when deploying new and potentially game-changing technologies. Member States, in collaboration with the multilateral system, could arguably do more in this regard. Typically, the initiative for this type of work comes from international organizations, think tanks, and non-profits.

One international organization pushing forward work in this area is the Organisation for Economic Cooperation and Development (OECD). Through a series of recent publications, the OECD has invited reflection on the uses of AI in the public sector to shed light on the policy concerns accompanying these technologies.[37] Among think tanks, the Carnegie Council for Ethics in International Affairs' Artificial Intelligence and Equality Initiative (AIEI) has been raising important questions about the potential for "AI systems [to] exacerbate structural inequalities."[38] JustPeace Labs, a non-profit, has broached the matter in a document entitled *Ethical Guidelines for PeaceTech* and other publications. They argue for the need to protect communities

Figure 3.

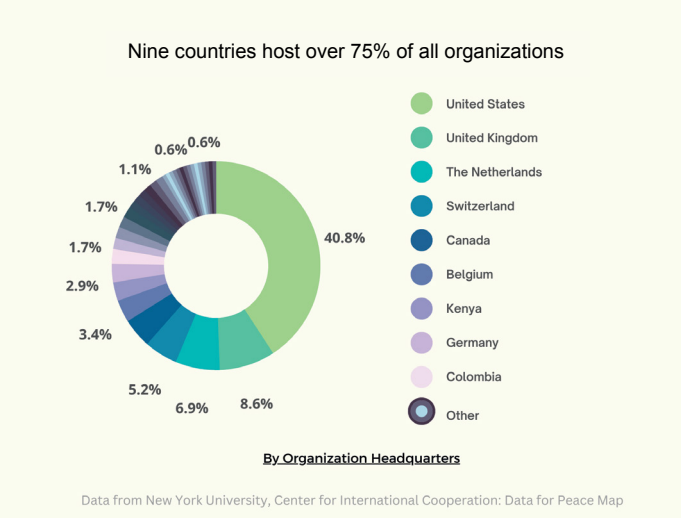

Figure 4.

on the receiving end of the technologies from "the risk of physical harm, shaming, retribution and group harms".[39] Furthermore, they emphasize the importance of deploying technologies hand in hand with local communities in order to create "effective partnerships for peace and human security".[40]

35 "Ecosystem Map: Data for Peacebuilding and Prevention," Center for International Cooperation, last accessed on 30 January 2023, https://cic.nyu.edu/data-for-peace-map.

36 Branka Panic and Paige Arthur, *Towards A Prevention and Peacebuilding Data Hub: Scoping the Future of Data Services and Capacity Building* (New York: NYU Center on International Cooperation, 2022).

37 Jamie Berryhill, Kévin Kok Heang, Rob Clogher & Keegan McBride, *Hello, World: Artificial Intelligence and its Use in the Public Sector*, OECD Working Papers on Public Governance No. 36 (Paris: OECD, 2019); and *OECD Public Governance Directorate, Governance Responses to Disinformation: How Open Government Principles can Inform Policy Options* (Paris: OECD, 2020).

38 "Artificial Intelligence and Equality Initiative," Carnegie Council for Ethics in International Affairs, last accessed on 30 January 2023, https://www.carnegiecouncil.org/initiatives-issues/artificial-intelligence-and-equality.

39 JustPeace Labs, *Ethical Guidelines for PeaceTech* (JustPeace Labs, 2017).

40 JustPeaceLabs, *Technology in Fragile Contexts: Engagement, Partnerships, and Positive Action* (JustPeace Labs, 2021); for a review of the importance

UN institutions, too, are increasingly sensitized to these broader reflections. In a recent briefing to the Security Council, the Under-Secretary for Political and Peacebuilding Affairs expressed several policy-related concerns toward the increased use of digital technologies. She brought up, for example, the possibility of automated systems being able to make decisions that impact human lives without humans being directly involved and highlighted the need for a "global digital compact" – as called for in the Secretary-General's *Our Common Agenda* – to outline principles for an "open, free and secure digital future for all". The United Nations, she said, "has a critical opportunity to build consensus on how digital technologies can be used for the good of people and the planet" and stressed that "collective action by Member States remains essential towards this goal".[41]

What would an "open, free and secure digital future for all" look like in the specific case of predictive technologies in peacekeeping? What are the challenges to that? Below is one example of the practical conundrum at hand:

> Moving toward a more automated predictive event analysis system would need to address cognitive and default biases already present in peacekeeping data analysis event taxonomies. For example, a decision to reduce complexity in the MONUSCO SAGE data entry form saw a large number of Mayi-Mayi groups collapsed into a single category for event perpetrator attribution, risking the identification of linkages where none exist. Ongoing taxonomy debates have also highlighted the particular challenges of using value- and/or legally-laden terms, such as 'terrorism' to describe events in a culturally and politically diverse analytical environment.[42]

Addressing this challenge is necessary if we are to take seriously the spirit of *Our Common Agenda* and the views of those communities most likely to be on the receiving end of decisions informed by the predictive models' output. It is a mistake to assume that all stakeholders share the same definitions, categories, and parameters which are fundamental to how the technology is calibrated. Put simply, these assumptions directly impact the kinds of predictions made. The slightest shift in taxonomy can change who the model predicts is likely to commit violence in the future. Building consensus around the parameters used to calibrate the models, and ensuring that the process is transparent and inclusive, will therefore be hugely consequential to their uptake within a multilateral system, which by definition is culturally and politically diverse. It will also help dispel the scepticism less powerful States may understandably harbour toward AI-based foresight capacity generated via funding from more powerful ones.

This represents a unique opportunity for Member States. By working with the multilateral system to advance shared calibration parameters and policies for achieving consensus inclusively, Member States may speed up the adoption of predictive technologies among UN peace operators. Not doing so will slow their adoption and risk foregoing the security benefits that may have otherwise accrued. The above is just one example of the challenges ahead, but it shows how important it is to have a shared understanding of the ethics involved. Indeed, there are several ethical aspects that should not be underplayed if Member States want to take these technologies forward. The following section will review the most salient.

of an approach inclusive of all stakeholders see Diana Dajer, "Cracking the Code of Tech for Peace: International Perspectives of Peacetech Research and Practice," *Reflections on Building Inclusive and Sustainable Peace* ed. Christine Wilson (London: British Council, 2018), pp. 70–79.

41 United Nations "*Political Affairs Chief Spells Out Double-edged Nature of Digital Technologies, in Briefing to Security Council*," 23 May 2022.

42 Dirk Druet, *Thematic Research Paper for the DPO Peacekeeping Technology Strategy: Enhancing the Use of Digital Technology for Integrated Situational Awareness and Peacekeeping-Intelligence* (New York: United Nations Department of Peace Operations, 2021), p. 16.

4. Uncertainty and the Precautionary Principle in Predictive Approaches to Peace

In 2018 the UN Secretary-General laid out a roadmap for a "digital revolution throughout the UN system".[43] The Departments of Peace Operations (DPO), Operational Support (DOS), and Management Strategy, Policy, and Compliance (DMSPC) subsequently produced a joint *Strategy for the Digital Transformation of UN Peacekeeping*. In it, they recognize that the vast quantity of information now available through digital technologies can play a role in conflict prevention – but that this is "accompanied by a high risk of collective data harms". According to the document, these may include abuses such as "breaches of confidentiality, behavioural surveillance, information disorder, information infrastructure sabotage or disruption". The authors also acknowledge that there are ethical questions around "data ownership, sovereignty and consent, social justice and potential social harm, as well as gender, race or other biases in algorithms for processing and analysing data".[44]

It is difficult to scratch the surface of the ethical challenges involved without being drawn into a rabbit hole of sorts: What are the legal and political risks of foreknowledge to civil servants and institutions? How do these technologies affect relations between major donor nations (that can afford them) and conflict-affected economies (that cannot)? What confidential information would institutional users of predictive software be handing over to private sector creators/owners (for example, via a hidden 'back door') by the very act of using it in a certain way? Could nefarious actors hack the tools to access private information or influence the predictions? How comfortable are we with humans being affected by decisions that are not entirely made by humans? Taken together, these organizational risks (see Figure 5 for a summary) underscore the "importance of choosing a do-no-harm approach" when using predictive technologies for peace operations.[45]

What exactly would a do-no-harm approach look like? How does an organization know that it is, in fact, not harming anyone? On one level, the answers to these questions depend on the entity using the technology. Predictive technologies possess a profound capacity for conflict prevention but can be damaging in the wrong hands. Indeed, the same tools that can be used for peacemaking can be used by non-state actors or illiberal governments for "social surveillance and control, repression, and racial profiling".[46] Massive data sets encompassing biometric, financial, geolocation, and behavioural information on entire populations introduce "new opportunities for authoritarian states or violent nonstate actors to control populations" and threaten "political participation, peaceful assembly, and freedom of movement".[47]

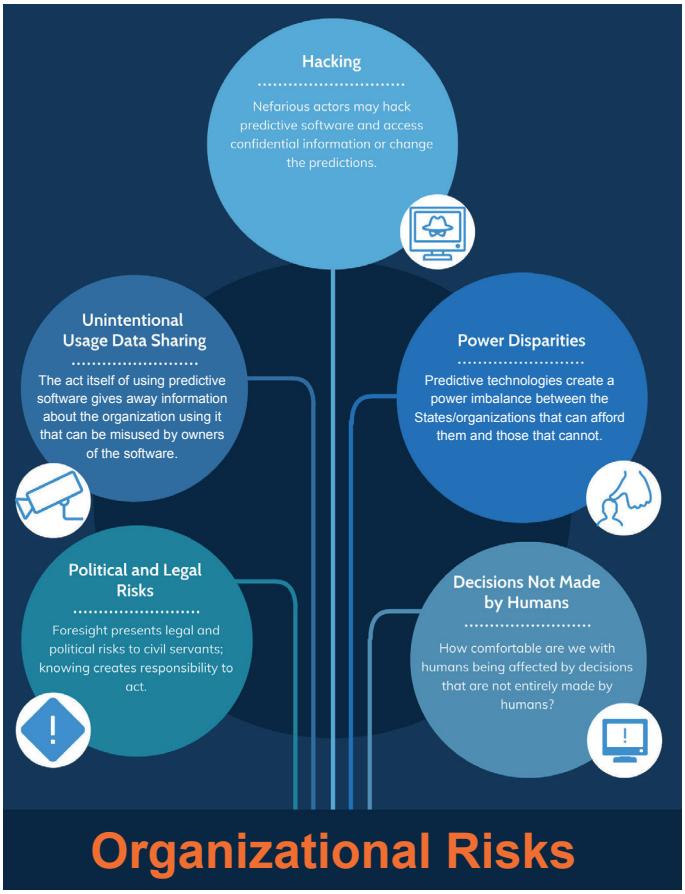

Figure 5.

43 UNDPO, "Strategy for the Digital Transformation of UN Peacekeeping" (New York, 2021), p 4.
44 Ibid., pp. 7, 13
45 Ibid., p. 12.
46 Eleonore Pauwels, *Counterterrorism and Violence Prevention: Safeguarding Against the Misuse and Abuse of Artificial Intelligence* (Washington, DC: Global Center on Cooperative Security, 2022), pp. 2–4.
47 Ibid.

The Center for International Governance Innovation similarly argues that policymakers, in cooperation with the private sector, must identify the areas in need of governance frameworks to prevent "behavioural nudging… and foreign adversarial influence on everything from elections to societal cohesion".[48] Multilateral organizations such as the UN could contribute significantly in this regard. Allan Dafoe, president and founder of the Center for the Governance of AI, states that "multilateral organisations could play a pivotal role in AI governance by providing a joint forum for the formulation, coordination, and dissemination of the cooperative norms between actors, enabling participating parties to signal sincere commitment to beneficial and shared AI development".[49]

While we must not underestimate the importance of "algorithmic transparency"[50] and actor usages that are inimical to democracy, on another level it may be the technology itself, irrespective of who is using it, that causes harm. Rainer Mühlhoff, a research associate at the Technical University of Berlin, raises the following interesting argument. Predictive analytics is used to forecast future behaviour of a target group or individual, Mühlhoff argues, but this prediction results from an analysis of enormous sets of behavioural data – the vast majority of which is about others. In other words, predictions about specific individuals are made "based on the data many unrelated individuals provided".[51] These predictions then inform differential treatment of that person or persons. In a peacekeeping or policing setting, it would look something like this: A specific person or group is predicted to have a higher likelihood of engaging in future violence, triggering closer monitoring of their activities. That prediction, however, is built on statistical modelling of data collected almost entirely about other peoples' activities. Mühlhoff calls this the "prediction gap" and discusses how it "challenges ethical principles such as human dignity and the (liberal) notion of individual privacy".[52]

In sum, "we face situations where an individual's (or group's) privacy is violated using data other individuals provide about themselves, possibly even anonymously." In this way, by establishing parameters of normalcy and deviance, "predictive systems produce and stabilise precisely the kinds of social differences and inequalities that they claim to merely detect in the world".[53] Riddles such as these have prompted scholarly conversations around the need to redefine the very notion of privacy. For example, given the scenario described above, it may no longer be possible to safeguard individual privacy simply via anonymization or other 'blurring' techniques (as it is now possible to know something about person A by crunching data from persons X, Y, and Z). A related conundrum is the matter of group privacy, as categories of people are increasingly targeted by algorithmic classification. Other scholars have debated the need to start thinking about protecting the privacy of groups as a central consequence of emerging technologies. They clarify that it should be considered "as an enhancement and safeguard for the individual right to privacy, rather than as a potential substitute for it".[54]

Another ethical concern is that predictive technologies engender trade-offs between action in the present and benefits in the long term. Presumably, the value of a prediction is that it can be acted upon now in order to influence events in the future. This is a tough trade-off because it involves comparing apples to oranges. The actions taken in the present are real, while the benefits in the future are hypothetical. The further in the future the predictions, the more this ethical concern is magnified, and the wider the prediction gap highlighted above. Consider the difficulty of intervening today toward certain groups based on what statistical models – that crunch data on mostly other groups – say is likely to happen in the far future. It is a cost-benefit analysis where the costs are selectively applied, immediate, and quite real, while the benefits are generic, far-off, and only projected. We may call this a temporal gap.

48 Meg King and Aaron Shull, "Introduction: How Can Policy Makers Predict the Unpredictable?," Centre for International Governance Innovation, 9 November 2020, https://www.cigionline.org/articles/introduction-how-can-policy-makers-predict-unpredictable/.

49 Allan Dafoe and Journal of International Affairs, "Global Politics and the Governance of Artificial Intelligence," *Journal of International Affairs*, Vol. 72 Issue 1 (2019): 121-126.

50 Robert Mazzolin, "Artificial Intelligence and Keeping Humans 'in the Loop,'" Centre for International Governance Innovation, 23 November 2020, https://www.cigionline.org/articles/artificial-intelligence-and-keeping-humans-loop/.

51 Rainer Mühlhoff, "Predictive Privacy: Towards an Applied Ethics of Data Analytics," *Ethics and Information Technology* Vol. 23 (2021): 675–690, pp. 675, 678.

52 Ibid.

53 Ibid.

54 Linnet Taylor, Luciano Floridi, and Bart van der Sloot, "Conclusion: What Do We Know About Group Privacy?" in *Group Privacy: New Challenges of Data Technologies* eds. Linnet Taylor, Luciano Floridi, and Bart van der Sloot (Springer Cham, 2017).

This cost-benefit analysis is necessarily political in addition to being purely statistical. As such, it is vulnerable to the possibility of manipulation by interest groups. Anja Kaspersen and Wendell Wallach, senior fellows at the Carnegie Council for Ethics in International Affairs, reflect on this risk in their critique of the idea known as "long-termism." The idea, popularised by philosopher William MacAskill, is "that the fate of humanity should be our top moral priority," and posits that the current generation (around 8 billion) should be making sacrifices to avert existential threats to future generations (hundreds of billions).[55]

MacAskill's long-termism is a common theme in the 2021 UN Secretary-General's manifesto, *Our Common Agenda*. The document urges that "now is the time to think for the long term, to deliver more for young people and succeeding generations and to be better prepared for the challenges ahead," and laments that "our dominant political and economic incentives remain weighted heavily in favour of the short term and status quo, prioritising immediate gains at the expense of longer-term human and planetary well-being". To rectify this, the manifesto proposes establishing a Futures Laboratory to "support States, subnational authorities and others to build capacity and exchange good practices to enhance long-termism, forward action, and adaptability".[56]

On some topics, like climate change, few would disagree with this stance, but on other matters the trade-offs are not so clear-cut. Kaspersen and Wallach fear that "legitimate concerns can easily be distorted and conflated with personal desires, goals, messianic convictions, and the promotion of deeply embedded political agendas and corporate interests". Investors in certain industries may divert attention from any short-term harm they cause by claiming future benefits outweigh the costs. Through this mechanism, "the well-intentioned philosophy of long-termism… risks becoming a Trojan horse for the vested interests of a select few". Or worse yet, it could give credence to the agendas of "technological elites pushing the development of technologies that have clearly demonstrated the potential to exacerbate inequalities and harm the wider public interest" by arguing that the long-term benefits outweigh these harms.[57] While much of the criticism of long-termism may be based on a distortion of its actual intentions, the need for an ethics of technological growth is clear.

The long-term benefits are often calculated – and argued for – using predictive statistical models. The same models also prescribe what to do today to change the course of future events: what costs to mete out and to whom. At a minimum, affected groups may want to know what data goes into the modelling and how predictions are made, but they may also want a voice in the debates around trade-offs and how they are defined. As digital data accumulates and machine learning capacity continues to compound (and as predictive technologies are applied to more levels of governance and deployed in more domains of everyday life), these calls for participation may increase. We may also anticipate these calls becoming a bigger part of definitions of open governance, civil society, and transparency. One UN Innovation Cell member expresses this spirit when pointing to the importance of "democratising foresight" and the need to "question whose vision of the future are we exploring, testing, and working towards".[58]

This scenario is further complicated by the fact that predictive models are based on deterministic assumptions about the nature of social phenomena. Confidence in a prediction is necessarily rooted in an understanding of events as determined by external causes. From a model's perspective, events, including human actions, are causally linked. This leaves little room for uncertainty, human agency, and emergent phenomena – all things that persist in the world and cannot be so easily measured. There is, therefore, a chasm between the deterministic approach of predictive models and the chaotic reality of the world fuelled by the fundamentally unpredictable nature of individual human agency. We may call this the uncertainty gap.

55 Anja Kaspersen and Wendell Wallach, "Long-termism: An Ethical Trojan Horse," Carnegie Council for Ethics in International Affairs, 29 September 2022, https://www.carnegiecouncil.org/media/article/long-termism-ethical-trojan-horse.

56 United Nations, *Our Common Agenda: Report of the Secretary-General* (New York: United Nations, 2021).

57 Anja Kaspersen and Wendell Wallach, "Long-termism: An Ethical Trojan Horse," Carnegie Council for Ethics in International Affairs, 29 September 2022, https://www.carnegiecouncil.org/media/article/long-termism-ethical-trojan-horse.

58 MinJi Song, "What if Uncertainty is the Path to Peace?" *Futuring Peace*, 8 October 2021, https://medium.com/futuring-peace/what-if-uncertainty-is-the-path-to-peace-29cfc5cd03d4.

5. Adopting a Precautionary Approach when Utilizing 'Peacetech'

All these organizational risks and logical gaps (see Figures 5 and 6 for an overview) strongly invite us to adopt the precautionary principle when considering the use of predictive technologies in peacekeeping. What would this precautionary approach look like? Member States, working alongside the UN policy research community and peace operators in the field, should spearhead work exploring this question.

Ideally, a precautionary approach would collaboratively identify and mitigate harms resulting from privacy breaches, be culturally and politically inclusive when calibrating models, and satisfy broader citizen questions around the prediction, temporal, and uncertainty gaps inherent in the technology. Models could be trialled in safe and controlled environments – in the UN Futures Lab, or DPPA/DPO, for example – before attaining 'approval' for wider use in a process similar to that used for medical innovations. Doing so would allow peacekeepers to garner the confidence required to use these technologies in practice. Not doing so may see practitioners either default to an excess of caution (and risk forgoing the peacekeeping benefits that may otherwise accrue) or possibly deploy the technologies in unintentionally harmful ways.

Collaboration with academia has yielded tools that can accurately predict conflict; it is now time to sort out common procedures and policies so that peacekeepers may adopt them and unlock their full potential. This is the central recommendation we hope may inform conversations on 'Peacetech,' including those around the High-Level Advisory Board on Effective Multilateralism and the upcoming Summit of the Future, in particular its New Agenda for Peace track.

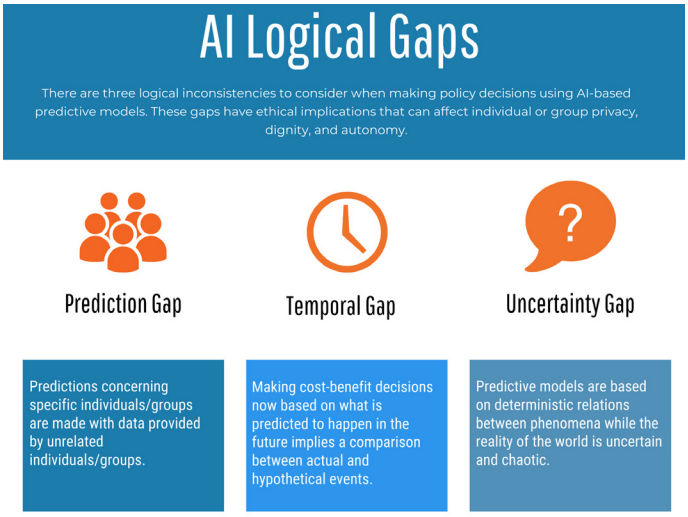

Figure 6.

United Nations University
Centre for Policy Research

About UNU-CPR

United Nations University Centre for Policy Research (UNU-CPR) is a think tank within the United Nations that carries out policy-focused research on issues of strategic interest and importance to the UN and its Member States. The Centre prioritizes urgent policy needs requiring innovative, practical solutions oriented toward immediate implementation.

The Centre offers deep knowledge of the multilateral system and an extensive network of partners in and outside of the United Nations. The United Nations University Charter, formally adopted by the General Assembly in 1973, endows the Centre with academic independence, which ensures that its research is impartial and grounded in an objective assessment of policy and practice.

cpr.unu.edu

New York (Headquarters)
767 Third Avenue 35B
New York, NY 10017
United States
Tel: +1-646-905-5225
Email: comms-cpr@unu.edu

Geneva
Maison de la Paix
Chemin Eugène-Rigot 2E
Geneva, Switzerland
Tel: +1-917-225-0199
Email: comms-cpr@unu.edu