July 2019

Ministry of Foreign Affairs of the Netherlands

DPPA
Preventing Conflict. Sustaining Peace

NYU | CENTER ON
—— | INTERNATIONAL
CIC | COOPERATION

DATA FOR PEACE & SECURITY

Report of the Practitioners Workshop on Harvesting Best Practices and Building a Community of Practice

About the Center on International Cooperation

The Center on International Cooperation is a non-profit research center housed at New York University. Our vision is to advance effective multilateral action to prevent crises and build peace, justice, and inclusion.

Table of contents

Summary and recommendations

As the world faces a significant upward trend in conflict—including a tripling of civil wars since 2007 and conflict increasingly prevalent in middle-income countries—practitioners in peace and security have sought to expand their toolkits to take advantage of the revolution in information gathering, data analytics, ICTs, and machine learning. A range of actors—multilateral organizations like the United Nations, governments, non-governmental organizations (NGOs), and the private sector—continue to invest in the capacity to make better use of data to promote peace and security.

In this context, there are many initiatives ongoing, often working in isolation. These include developing new practical and innovative data sources, better ways to mine and analyze data, algorithms, use of technologies like blockchain, and data-driven applications. New approaches have been developed with the objective to better define where the risks and threats to peace and security are most urgent, and where our assistance is needed most, in order to inform decision-makers with better insights and information, and to generate a greater impact on the ground. In many cases, these efforts result in specific user-products which in turn inform the workflows of policy advisors, advocacy officers, conflict prevention, peacebuilding and development practitioners, and researchers around the world.

Nonetheless, the peace and security field lags behind other sectors, like the humanitarian and development fields, in using data to drive innovation and greater impact. It also lacks a community of practice across sectors (government, private sector, multilaterals, civil society, academia). This community is needed in order to identify and build on lessons learned, prioritize the most important needs and gaps, spark further innovation, prevent duplication, and to make sure product development is demand-driven.

In this workshop, which took place on March 20, 2019, in New York City, participants discussed how innovations in data are transforming the methods and the effectiveness of those working on early warning, conflict prevention, peacebuilding, stabilization, and international security. It was co-organized by the Netherlands Ministry of Foreign Affairs, the NYU Center on International Cooperation, and the UN Peacebuilding Support Office.

The objective of the workshop was to generate ideas that will allow for a better use of data to promote peace and security by learning from each other, building on existing experiences, and sparking new partnerships. The workshop aimed to:

- Advocate for and showcase innovative data solutions to challenges in peace and security sector

- Create an international multi-stakeholder community of practice around data specifically in the peace and security field

There were more than 70 participants who came from around the world, including Australia, Nigeria, Ukraine, Kenya, Germany, the Netherlands, the United Kingdom, and Canada. They represented a broad set of sectors, including governments and multilaterals, international and civil society organizations, think tanks, and the private sector.

The day was organized around a mix of quick-and-dirty thematic sessions called "sprints" (which aimed to collect and exchange as much information and insights as possible from specific projects), a lunchtime marketplace of ideas, and plenary feedback and discussion. A total of 28 projects were showcased during the day: 17 in the sprints and 11 in the marketplace of ideas.

A post-workshop evaluation suggested a positive response. Out of 13 respondents, all but one said that the format was "just right" in terms of flow and dynamism, and quality scores for sprints, panels, etc., were mainly at the high end of the scale (85 percent of scores of "good" or "very good"). Every respondent said they had made at least one new professional contact, but most (85 percent) made at least 3, and a good number (38 percent) made 5 or more new contacts. As one participant summed it up: "Great experience to share what are data challenges and see what others are doing as well as sharing and networking."

The workshop yielded **five insights and opportunities**:

- We are on the right track in initial attempts to use data-driven approaches.
- There is a role for a community of practice.
- Sharing data, applications, and platforms makes the whole greater than the sum of its parts.
- There is a need to focus on building capacity.
- Communicate, communicate, communicate!

And **four challenge areas**:

1. Recognize that there are different kinds of expertise.
2. Get real about bias in data and communicating about this bias.
3. Find ways to reduce political sensitivity of findings.
4. Improve our own practice

From this the organizers settled on the following **recommendations**:

Recommendations

1. **Establish and support a community of practice on data for peace and security:**

   1.1. Offer continued dialogue among practitioners on data for peace and security. Disseminate calls for proposals, events and studies and facilitate meet-ups at key events over a LinkedIn Group.

   1.2. Capture and share lessons learned on both success <u>and</u> failures of initiatives. Cultivate buy-in on data for peace and security at senior level to encourage risk taking by showcasing ideas and practices of international organizations, governments, civil society organizations and academic institutions. Consider development of automated data-sharing platforms.

   1.3. Priority thematic areas to explore include early warning and risk modelling; use of text mining applications; ethics of artificial intelligence in peace and security applications; innovative approaches to monitoring, evaluation, and impact measurement.

2. **Build and sustain capacities for data driven approaches among CSOs, multilateral organizations, and government:**

   2.1. Enable donors to support initiatives from the Global South and initiatives related to South-South and triangular cooperation through a mapping exercise. The exercise should introduce initiatives to new audiences and help initiatives from the Global South disseminate their findings.

   2.2. Further invest in early warning and ensure early warning risk assessments are connected to response mechanism. Connecting existing datasets and systems is a 'low-hanging fruit' for drastically improving data driven assessment and planning in a cost-efficient manner. Establish a practice of communicating data to decision makers.

   2.3. Bring machine learning and big data to bear on enhancing integrated multidimensional risk analysis—promoting an understanding of how disparate risk areas are interlinked to provide better policy advice regarding prioritization and sequencing of actions.

   2.4. Consider (through a follow-on workshop) the inclusion of data for peace and security as part and parcel of achieving existing policy objectives within your organization.

   2.5. Work with local actors to identify capacity gaps and political blockages for data-driven approaches; to raise their visibility with governments, interested NGOs, and private sector actors; and to create new partnerships with external actors like private sector, social entrepreneurs, and "data for good" clinics.

3. **Improve design and use cases:**

   3.1. Focus on approaches with up-front investments in design and ensure user-friendly products (through end-user analysis) that support effective communication.

   3.2. Encourage private sector partnerships within the community of practice on data for peace and security; map pro bono data clinics in the private sector with of relevance to peace and security issues.

Opening discussion

The opening panel, moderated by Sheldon Himelfarb of PeaceTech Lab, kicked off with a discussion of some of the most important challenges and opportunities in the field today. Panelists agreed that data is radically transforming the way we work—with Robert Kirkpatrick of UN Global Pulse predicting that it would be "wildly disruptive" and generate new social sciences. They registered excitement around the possibility of using new data sources to understand social cohesion, as well as the gains in efficiency and cost in collecting for collecting new data that they were

Sheldon Himelfarb of PeaceTech Lab leads the opening panel.

already seeing. Advances have already been made around measuring things like peacebuilding, the role of peace operations, and the impact of institutions such as the justice sector in a peacebuilding process. Given the focus on peace and security, panelists like Rachel Brown of Over Zero and Monica Nthiga of Ushahidi highlighted the opportunities that crowdsourcing and other kinds of community-driven data collection can have for strengthening human relationships and generating new tools for empowerment of ordinary people.

They also agreed on key challenges, many summarized by Patrick Vinck of the Harvard Humanitarian Initiative. These include lack of access to existing data, dealing adequately with biases in data, and the need for more data literacy (and "data advocates") among policy actors. Importantly, there was agreement on the challenge of ensuring adequate respect for local expertise and perspectives, and for the physical safety of those collecting or sharing sensitive information.

Noting that the **private sector** is light years ahead in leveraging machine learning, AI, and data analytic tools for their businesses, Himelfarb suggested that there was a lot to learn from the private sector's more competitive approach, asking if what we need is more competition rather than communities of practice. Panelists had mixed views. Vinck agreed that the private sector is ahead and the challenge is to catch up—with the legal framework being farthest behind on issues like privacy. Brown observed that there is plenty of competition in the NGO sector and that it is important to strengthen accountability to communities rather than increase competition for donor funding. For Kirkpatrick, competition in the tech sector is not always virtuous; for example, it has massively accelerated polarization. This said, he predicted that there will be ample commercial opportunities, since at the end of the day, conflict is bad for business. We are already in a blizzard of information, panelists observed, and more is coming. As the fields of network science, AI, and machine learning evolve, opportunities and risks will only multiply.

The sprints

The day was organized around dynamic, hour-long "sprints" that showcased practical solutions to problems that practitioners face in the peace and security field.

Sprint 1: Dealing with complexity

The first sprint introduced leading-edge approaches to understanding and analyzing multidimensional risks and how they interact and influence peace and security, as well as how predictive analytics can inform our risk models. Innovations reviewed in this sprint were: **ECOWARN**, the early warning system of the Economic Community of West African States; the US State Department's **Instability Monitoring and Analysis Platform**; the Hague Centre for Strategic Studies' (HCSS) **Conflict Risk Assessment Project**; and the **Peace Perceptions Poll**, which is a collaboration of RIWI Inc. and International Alert.

Practitioners in the peace and security field are used to analyzing complex systems, and new data-driven approaches open up both new fields of knowledge and additional layers of complexity. One common thread concerned how these systems deal with **data bias**—whether they are relying mainly on their own data (ECOWARN), or drawing from other sources such as the Armed Conflict Location and Event Data Project (ACLED). ECOWARN discussed its capacity-building approach with its country-level monitors as well as the triangulation that it does to reduce the bias in its data, through field missions, media analysis, and other methods. In the Peace Perceptions Poll discussion, a productive debate took place on incorporating voices from more local contexts, and expanding the analysis from a country level comparison to diving into regional nuances. Presenters for the Instability Monitoring Platform discussed some of the data-driven tools and communication techniques they use to help policymakers understand conflict and crisis issues.

Another thread surfaced around capturing multidimensional risks at sub-national level. HCSS's Conflict Risk Assessment Project, for example, has a sub-national focus that makes it more sensitive to conflict indicators particular to a region, group, or community. Nonetheless, discussion focused on ways to strengthen this focus to capture conflict dynamics that transcend these lines of demarcation, including getting even more localized inputs.

Complexity also goes hand-in-hand with **context sensitivity**, and all of the presentations addressed they ways in which their applications tried to address this issue. Discussions focused on issues such as how to take into account how social cohesion would be measured differently in different countries (e.g., Colombia versus the Democratic Republic of the Congo), through nationally tailored questions. Another example concerned how to collect data effectively from sub-national groups, but still harmonize it with data from the national and international levels. Cutting across the discussions was a recognition that dealing with complexity requires

interpretation—introducing a qualitative element, in some cases local knowledge—rather than just algorithms.

Finally, the human element of dealing with complexity was discussed, including the need for **thoughtful design around use-cases** so that complexity does not overwhelm end users. For example, ECOWARN has a visualization system tailored to decision-makers called the Regional Geospatial Portal; using the portal, it became clear that the greatest human security threat in the region is car accidents—not terrorism, where many resources have been concentrated.

Sprint 2: Addressing polarization

In the second sprint, participants addressed data innovations to deal with radicalization, deep fakes, hate speech, misinformation and disinformation, and similar phenomena. Applications included in this sprint included **Truepic**, which uses blockchain to verify photo and video content; **Qatalog**, an online platform for monitoring social media and public radio broadcasts, presented by UN Global Pulse; the **Sentiment Analysis and Digital Focus Groups Project** of the UN Department of Political and Peacebuilding Affairs; and a project on collecting high-resolution **"ground truth" data for countering violent extremism** by Elva Community Engagement.

Increasing polarization of societies is affecting the work of peace and security practitioners globally, and **finding both "upstream" and "downstream" approaches to tackling it is critical**. The applications showcased intervene at various moments of polarization. For example, Elva's approach combines upstream approaches, by detecting and countering extremist influence early, and downstream approaches in which, for example, they try to disrupt violent extremists' income streams. Truepic's approach is more downstream, focusing on trying to shift users' perceptions of polarizing content as it circulates. Both of the UN-led applications (Qatalog and the Sentiment Analysis project) provide real-time information on perceptions that could be useful to understanding how polarization is influencing political violence, including around important moments such as elections. In particular, the Sentiment Analysis project aimed to understand how (and by whom) political narratives are shaped.

Participants noted that that **polarization requires a whole of society approach**, and no data-driven innovation is a panacea. For example, with Truepic, it was noted that political use of fake images is nothing new; what is new is the easy transmission and circulation within larger groups of people that previously were disconnected, as well as lax regulatory frameworks. Truepic's advance therefore provides an important practical solution to verify images, but there is still a need for larger solutions to complement it.

Participants take a closer look at the German Federal Foreign Office's data analysis tool.

Across the four discussions, there was a common theme around the **technical difficulty of using data-driven approaches to detect polarizing speech or images**. The applications used a wide range of methodologies, from natural language processing to blockchain to household surveys and qualitative approaches (key informant interviews, focus groups)—reflective of the fact that understanding and addressing a phenomenon as complex as polarization cannot be done through single approaches or a "magic bullet." As examples of challenges, it was noted that in Arabic, a lot of speech uses sarcasm, which is challenging for natural language processing. There was also a discussion about the recent massacre in Christchurch, in which the extremist livestreamed the bloodshed on Facebook for more than an hour before being detected.

Discussants also emphasized that people whose intent is to drive polarization and extreme views are quick to adapt to new technological means to detect and prevent them. Some even questioned what would prevent someone committing an extremist act to use a platform like Truepic to "prove" that the act was real—with the response that nothing could prevent it, but of course, in so doing, this would also provide a rock-solid piece of evidence that could be used to prosecute them. The need to "do no harm" in polarized environments was also raised in the discussions. In the presentation from Elva, for example, participants spoke of the steps that can be taken so that informants are kept safe. These include constant security assessments, anonymization, limiting sensitive questions, and not sharing specific data with local authorities.

Sprint 3: Communicating for decision-making

The third sprint discussed innovative uses of visualization and other communication tools to connect decision-makers with timely and relevant data sources, including in peacekeeping operations, HQ situation rooms, and defining the costs of inaction. Innovations included **PREVIEW**, the German Federal Foreign Office's data analysis tool for crisis early warning; **Whiteflag**, presented by the Royal Netherlands Air Force and the International Committee of the Red Cross; the UN Development Programme's **Crisis risk dashboard for early warning** and new uses of **data and analytics for Security Council briefings** from the UN Department of Peace Operations; and the US Holocaust Memorial Museum's **Early Warning Project**.

The applications were designed to meet a wide range of communications needs for decision-makers in the peace and security field, with a focus on **situational awareness and early warning** for outbreaks of violence or atrocity crimes. There was an accent on being able to prioritize both at national and sub-national level the highest risks, so that resources could be quickly deployed to mitigate them. For example, data and analytics from the UN Department of Peace Operations could be used in Central African Republic to ensure that deployment of peacekeeping troops match the highest level of need for internally displaced persons, or in Democratic Republic of Congo to help placement of air assets. The Early Warning Project described its unique approach, which focused on identifying plausible scenarios or sets of events for large-scale attacks on civilians, rather than assessing whether such risks are "high" or "low."

A consistent line of questioning in the discussions surfaced around the extent to which these applications were driven by the demands of decision-makers, and if so how effectively they were **tailored to decision-makers' needs**. The discussion of the Early Warning Project turned to the

issue of how to ensure a focus on product design from the start, so that the data is digestible for targeted end-users. One of the UN presenters noted that there was an inverse relationship between granularity of data and level of decision making (i.e., higher levels of decision-making demand less granularity), while another developed the idea of a "kitchen" of data ingredients, and you "cook up" analytical products based on the decision-making need. Whiteflag had a unique approach among the presenters in its use of blockchain as a way to create a trusted, neutral platform valued by decision-makers in rapidly changing, politically volatile, fragile contexts. Whiteflag provides a means of secure, near real-time, and verified information exchange in conflict zones allowing, for example, neutral parties like NGOs or hospitals to make themselves known to all, or only to specific parties involved for deconfliction purposes. The blockchain protocol also supports other use-cases, such as marking danger zones, providing proof-of-life, and coordination of humanitarian assistance after a natural disaster. Critically, participants in this sprint agreed on the need for **senior leaders to be champions for data**.

Roudabeh Kishi leads a sprint on ACLED.

There was broad agreement that data-driven approaches had already done much to advance efforts to **monitor and evaluate outcomes** of peace-related initiatives in the field. But participants also noted a specific challenge when it came to requests to provide data on political progress, which is often at the heart of peace-related goals. For example, the UN presenters observed that it can be more challenging for them when they are asked to provide monitoring data on the more political Security Council mandates, such as the one in Kosovo.

Similar to other sprints, there was a strong focus on data bias issues. The challenge was framed differently, however, as it focused on **how to communicate about bias, gaps in data, and "dirty" data**. There was a widely shared concern that decision-makers often do not have enough **data literacy** to understand how best to interpret data they are seeing, and a concomitant reluctance on the part of data analysts to communicate about bias, for fear of making data seem useless for decision making.

Finally, there was a lot of talk on **the politics (and politicization of data) for decision-making,** which is an especially challenging issue for practitioners working in the heavily politicized peace and security space. Many expressed the view that statistics can help to depoliticize sensitive issues around risks, through providing an objective source. But it was also emphasized that **policymakers should be deciding on risks, rather than data scientists**. One reason is that data science can tell you only what data does and does not say; another is that many of the statistical models described are not causal, making it difficult to use them to develop recommendations. In relation to issues of severe government sensitivity toward "uncomfortable"

data like risks and triggers, participants advocated for finding ways to have private discussions with governments. One person described this as "early warning as persuasion," where one praises in public and criticizes in private.

Sprint 4: Improving practice and understanding impact

The fourth sprint focused on data-driven approaches and new data sources to create indicators for improving the effectiveness of peacebuilding practice that are useful for monitoring and evaluation of interventions such as real-time ICT applications, location-specific data, crowdsourced information, and closer interaction between humans and machines. Applications showcased were **Sentry** from Hala Systems, the **Armed Conflict Location and Event Data** Project (ACLED), the **Syria Mapping Project** from The Carter Center, and the Human Rights Data Analysis Group's efforts to **locate hidden graves in Mexico using machine learning**.

A critical part of practice discussed was the **importance of creating agency among conflict-affected peoples**. Sentry fosters this approach through ensuring engagement among the most marginal and disconnected populations in war zones, disaster areas, and wilderness. The Syria Mapping Project does so through working with a wide and deep set of stakeholders in the Syrian conflict to help shape the larger mediation process. The project on locating hidden graves in Mexico is designed to give families of the missing information that can be used to take action with prosecutors and other state authorities. These applications take into account the importance of human relationships in their approaches.

There were further discussions on the **centrality of product design and attention to end users** in improving practice. Sentry, which is in the process of developing a business model, led a discussion on how to design usable, practical tools. The Syria Mapping Project also highlighted its use-cases with Special Envoys and other mediators, as well as with humanitarian actors, emphasizing the importance of cross-sectoral collaboration. Participants agreed that greater attention to design is critical in creating momentum for uptake of data-driven approaches at the practitioner-level, with initiatives like these paving the way.

Each of the applications made a distinct contribution to discussions on how to improve practice in conflict-affected contexts. A crosscutting issue, particularly in relation to understanding impact, concerned ways to develop **neutral, trustworthy information.** There was a rich discussion in the session on ACLED, which has become one of the most widely used data sets (including by many of the applications at this workshop) because it has found a balance between flexibly adapting its taxonomy to emerging needs and taking a rigorous approach to managing the data itself. The presenter led an in-depth discussion on its approaches to bias, describing the work of their coders and how they make corrections. Like in other sprints, data bias was discussed across the four presentations, including a lively debate on event-size bias in the session on hidden graves in Mexico. It was noted that many of the applications in the peace and security field (including those presented at this workshop) rely on one data source or on a flawed data source, and thus there is a need for a cautious approach. As one participant noted, "When selection bias meets confirmation bias, they live happily ever after."

Marketplace of ideas

In addition to the four sprints in which 17 applications were presented, another 11 applications were showcased at the lunchtime marketplace of ideas. These included a fascinating variety of tools, including chatbots for people in crisis situations, early warning tools, indicators on UN peacekeeping, predictive models for regime change, and a tool to track flight patterns of those engaging in potential illicit flows. The participants were:

The World Bank presents its risk monitoring platform at the marketplace.

- Visualizing Climate and Violence Risks in the Sahel, Igarapé Institute
- Data for Risk Monitoring, The World Bank
- ChitChat: Communication with People in Crisis, Centre for Innovation, Leiden University
- VForecast: Predicting Adverse Regime Transitions, Predictive Heuristics
- Countering Violent Extremism through Lifestories, The Hague Center for Strategic Studies
- 44 Indicators on SDG16+ Actualization: A Framework for Policy Analysis, Institute for Economics & Peace
- UN "SAGE" Incidents/Events Tracking Database System, UN Operations and Crisis Centre/Peace Operations
- Protection of Civilians, PAX
- Early Warning Dashboard and Predictive Models, Ministry of Foreign Affairs of the Netherlands
- Visualizing Complex Conflict Environments Through Machine Vision and Open Flight Tracking, C4ADS
- Mathematical Models and Data Science Measurements of Sustainable Peace, Queens College

Conclusion: Opportunities and challenges

Data is already transforming our world, and the way that practitioners in the peace and security field work will be radically different 5, 10, and 15 years from now. But data is not a panacea; it can be leveraged for better or for worse. If it is going to help improve practice and outcomes, it must be used sensitively, and we must build capacity and intellectual infrastructure so that people and policymakers alike can grapple with it thoughtfully.

The opening panel provided many of the threads that then were pulled through the workshop. **Patrick Vinck** cautioned against misuse of data-driven approaches and the need to examine bias closely. **Robert Kirkpatrick** expressed a hope that data would be able to tell us more about social cohesion and how it operates; how social capital is created and destroyed. This thinking tied into many of the presentations, particularly on polarization. Some of these looked at push and pull factors for radicalization; others looked at using technology to reduce the potential for polarization, by verifying the truth of kinds of content (like images).

Monica Nthiga introduced the need to care with crowdsourced data, especially regarding hate speech but also beyond. Several projects discussed the potential for harms for the very people we should be trying to protect, and always the need for protective and mitigating measures for the courageous people who are most at risk. This has been a long-term theme for humanitarian sector, which we can do more to learn from. "Do no harm" is critical.

Rachel Brown emphasized that data is not a substitute for human infrastructure. We need to examine whether data collection will be used to build relationships or to extract information. This tied nicely to the need particularly from a peacebuilding and human security perspective, to do data collection in a way that contributes to building local relationships. ECOWARN's approach, working at the local level through civil society organizations and community-based organizations, is an example of this, but there are many others.

This framing helped to surface some **key insights and opportunities**:

1.   We are on the right track.

Data, tech, and analytics are already transforming our work to promote peace and security. But we are just getting started. There is much untapped potential. Mapping available data sources and potential applications is critical in making forward progress, given the large number of possible use cases discussed at the workshop, including: protecting civilians; increasing situational awareness of and facilitating timely early warning for risks to peace and security; increasing effectiveness of peace operations; countering hate speech, fake news, and misinformation; increasing agency of communities to shape conflict prevention and

peacebuilding policies and initiatives; measuring impact and monitoring and evaluating progress; informing and facilitating decision-making processes in the field of peace and security; and protecting human rights.

2. There is a clear role for a community of practice

Collective investment in a multi-stakeholder community of practice is needed to reap greater benefits, by making available to as many people as possible the knowledge already developed, lessons already learned, and use-cases already identified. The Data for Peace & Security LinkedIn Group that came out of this workshop is a modest step. Follow-up events would be welcomed as well as other ways to strengthen and expand the community of practitioners. It will be important to invest in reaching actors in the global North and South to include their expertise and experience into the global community of practitioners, as well as to showcase use-cases in all of these regions.

3. Sharing data, applications, and platforms makes the whole greater than the sum of its parts

Benefits can accrue from sharing (re-structured) data-sets, (parts of) data-driven products, algorithms, etc.—as well as opening up the products that we have developed by making them as open-source as possible (while of course respecting that due to privacy or security reasons this will not always be possible). We could invest in platforms where these capacities can be shared. Larger organizations with the means to invest in advanced data analytics can team up with smaller organizations to ensure their access as well.

4. There is a need to focus on building capacity

Collectively resourcing proven and/or promising organizations, initiatives, and data-driven products will continue the ongoing process of capacity building. This also calls for greater coordination to make sure resources flow to those places where demand from affected populations (and from practitioners seeking to promote their peace and security) is highest. It will be important to invest in data literacy and the right skillsets for employees within organizations and people within societies. There may be opportunities to develop joint funding vehicles where this coordination can take place as well. Also, through secondment of staff (e.g., data scientists), capacity can also be built and experiences shared.

5. Communicate, communicate, communicate

We need to communicate as much as possible about the potential and need for data, tech, and analytics to promote peace and security. We need to communicate our success in order to increase the support base for these innovative undertakings, and invest in outreach strategies about the potential benefits of data for peace and security. Senior leadership, in particular, will need to hear about this in order to gain support for increasing investments in this field.

But there was also consolidation around **four key challenge areas**:

1.   Recognize that there are different kinds of expertise

Technical expertise alone will not solve complex challenges faced by practitioners in the peace and security field. It is therefore important to identify and integrate many sources of expertise relevant to data-driven approaches to peace and security, from data scientists to local organizations and communities. It is also important that initiatives are demand-driven and that use-cases are clearly defined.

Participants highlighted the need to de-center our assumptions about expertise, and how we define it. Which group of people has the brainpower to make predictions by picking up on signals? Sometimes it is the specialists and sometimes it may be local communities. An additional question concerns who has the expertise to really understand the limitations of data, including making judgments about taking action and defining what is considered a risk. For developing data-driven policies and products it is important to put the user first and fine-tune products to user needs.

2.   Get real about bias and communicating about bias

Many of the contexts in which practitioners promote peace and security are "data poor" (in the traditional sense); often there is not enough data available and data-sources are limited. Therefore, it is important to navigate the tension between the limitations of data and the need to inform policymakers and create effective change. Users of data-driven applications need to be informed about the potential biases of data-sources fed into the applications. After all, the key rule of data-science still applies: garbage in = garbage out.

As one participant said, data is always lying to us, and we therefore need models to analyze bias, or else we will be captured in those lies. (But as another said: all models are wrong, but some are useful.) Machine learning can not only reproduce bias but also amplify it. Therefore, constant recognition of the bias and the ability to communicate effectively about it to decision-makers is critical. Yet participants recognized a bind here: there was a real tension between understanding that there is bias and not letting that completely prevent us from having an impact on decision making with data. At a minimum, participants argued that it is important to get additional information to decision makers (like Ambassadors and Special Envoys) that at least reduces their own personal bias.

3.   Find ways to reduce political sensitivity of findings

Issues of peace and security are inherently political, so it is critical to find ways to desensitize data-driven findings, so that they will be acknowledged and used effectively by decision-makers. Approaches seeking to assess risks of mass atrocity, or predict where mass graves might be, are very sensitive issues with the very actors—national governments and local authorities—that most need to be influenced. Whether what is at issue are grave sites in Mexico, mass atrocities in Democratic Republic of the Congo, or images coming from a conflict like Yemen or Syria, it can be

hard for findings to be "heard" by people able to make a difference. There is thus a need to find ways to use scientific, data-driven approaches in order to depoliticize issues. Statistics can provide a level of neutrality, since it is "science" telling authorities about risks, rather than an NGO or a foreign government.

4. Improve our own practice

Practitioners on the ground are facing complex challenges. If they have to wrestle with poor design or lack of a clear use case, they are less likely to engage with data-driven approaches. Therefore, engaging from the start with the practice of design, users, and metrics of success is important. A consistent question was how does this field get better at the practical stuff: design, UX, user-driven approaches, business models, and metrics for success? Solving these issues are also critical for successful communication. As one participant said: we crawl over the finish line and leave the retail business to the end. How do we know if these applications are successful? Can we focus on developing measures of user adoption or success?

Annex: Workshop agenda

9:00 –9:15 Registration and breakfast	Registration opens in front of Room 914 on the 9th floor of NYU Kimmel Center; light breakfast and welcome offered
9:15-9:30 Plenary	Introduction: • Bas Bijlsma - Senior Policy Officer at the Ministry of Foreign Affairs of Netherlands Welcoming remarks: • Ms. Sandra Pellegrom, Head of the Development, Human Rights and Humanitarian Department, Permanent Mission of the Kingdom of the Netherlands to the UN Plan for the day • Paige Arthur - Deputy Director of the Center on International Cooperation at NYU
9:30-10:15 Panel discussion	**The biggest challenges in data applications to peace and security: What are the big-picture, forward-looking trends in the field?** Moderator: Sheldon Himelfarb, President and CEO, PeaceTech Lab Interactive panelists: • Rachel Brown, Executive Director, Over Zero • Robert Kirkpatrick, Director, UN Global Pulse • Monica Nthiga, South to South Lab Manager, Ushahidi • Patrick Vinck, Director, Peace & Human Rights Data Program, Harvard Humanitarian Initiative
10:15-11:15 Sprint sessions #1	Sprint #1: **Dealing with complexity:** leading-edge approaches to understanding and analyzing multidimensional risks and how they interact and influence peace and security, as well as how predictive analytics can inform our risk models • **Room 903 - ECOWAS Warning and Response Network (ECOWARN)** by Kebba Touray, Economic Community of West African States (ECOWAS); Doug Bond and Sean Yeo, Harvard University and Virtual Research Associates, Inc. **Facilitator:** Henk-Jan Brinkman, UN Department of Political and Peacebuilding Affairs • **Room 905 - Instability Monitoring Analysis Platform** by Melissa Duell and Christopher G. Istrati, The Bureau of Conflict and Stabilization Operations, US State Department **Facilitator:** Michael Colaresi, University of Pittsburgh • **Room 907 – HCSS Conflict Risk Assessment Project** by Hannes Roos, The Hague Centre for Strategic Studies **Facilitator:** Peter Kirechu, C4ADS

	• **Room 909 - Peace Perceptions Poll** by Julian Egan from International Alert and Leah Moncada, RIWI **Facilitator:** Giovanna Kuele, Igarapé Institute
11:15–11:30	**Coffee break**
11:30-12:30 Sprint sessions #2	Sprint #2: **Addressing polarization:** data innovations to deal with radicalization, deep fakes, hate speech, misinformation and disinformation, and similar phenomena • **Room 903 - Collecting high-resolution ground truth data for countering violent extremism** by Jonne Catshoek and Mark van Embden Andres from Elva Community Engagement **Facilitator:** Patrick Vinck, Harvard Humanitarian Initiative • **Room 905 – Sentiment Analysis and Digital Focus Groups Project** by Daanish Masood and Martin Waehlisch, UN Department of Political and Peacebuilding Affairs **Facilitator:** Rina Amiri, Center on International Cooperation-NYU • **Room 907 - Qatalog: An Online Platform for Monitoring Social Media and Public Radio Broadcasts** by Jeremy Boy, UN Global Pulse **Facilitator:** Monica Nthiga, Ushahidi • **Room 909 – Photo and video verification using blockchain** by Mounir Ibrahim, TruePic **Facilitator:** Christopher Georgen, Topl
12:30-13:00 Report back from group sessions	Facilitators: Paige Arthur (CIC-NYU) and Ayham Al Maleh (UN Department of Political and Peacebuilding Affairs) **Rapid-fire report back from the facilitator of each group (2-3 minutes) and posting to the wall of ideas in plenary space.**
13:00-14:15 Walking lunch in the marketplace of ideas	Walking lunch & networking through the marketplace of ideas in plenary space, which will feature additional innovative applications and solutions. Featuring: **Visualizing Climate and Violence Risks in the Sahel** Giovanna Kuele, The Igarapé Institute **Data for Risk Monitoring** Sophia Armanski and Phoebe Girouard Spencer The World Bank **ChitChat: Communication with People in Crisis** Josje Spierings Centre for Innovation, Leiden University

	VForecast: Predicting Adverse Regime Transitions Andreas Beger, Predictive Heuristics **Countering Violent Extremism through Lifestories** Arlinda Rrustemi, The Hague Center for Strategic Studies **44 Indicators on SDG16+ Actualization: A Framework for Policy Analysis** David Hammond, Institute for Economics & Peace **UN "SAGE" Incidents/Events Tracking Database System** Rajkumar Cheney Krishnan from UN Operations and Crisis Centre/Peace Operations **Protection of Civilians** Hans Rouw, PAX **Early Warning Dashboard and Predictive Models** Bas Bijlsma, Ministry of Foreign Affairs of the Netherlands **Visualizing Complex Conflict Environments Through Machine Vision and Open Flight Tracking** Peter Kirechu, C4ADS **Mathematical Models and Data Science Measurements of Sustainable Peace** Larry Liebovitch, Queens College

14:15-15:15
Sprint sessions #3
	Sprint #3: **Communicating for decision-making:** innovative uses of visualization and other communication tools to connect decision-makers with timely and relevant data sources, including in peacekeeping operations, HQ situation rooms, and defining the costs of inaction
	
	• **Room 903 - PREVIEW, the Federal Foreign Office's data analysis tool for crisis early warning**
	by Nina Bergmann and Georg Kalckreuth from the German Federal Foreign Office
	    Facilitator: Doug Bond, Harvard University
	
	• **Room 905 – Whiteflag**
	by Timo Schless from Royal Netherlands Airforce and Vincent Graf from the International Committee of the Red Cross
	    Facilitator: Jeremy Boy, UN Global Pulse
	
	• **Room 907 – Data for decisions at the UN**
	1. Crisis risk dashboard for early warning
	by Corrado Scognamillo and Shouryadipta Sarkar from UNDP;
	2. Data, analytics, and Security Council briefings
	by Avishan Bodjnoud from UN Peace Operations;
	    Facilitator: Evan Cinq-Mars, CIVIC
	
	• **Room 909 – Early Warning Project**
	by Lawrence Woocher from United States Holocaust Memorial Museum
	    Facilitator: Miranda Sissons, Center on International Cooperation-NYU

15:15-15:30	**Coffee break**

15:30-16:30
Sprint sessions
#4
	Sprint #4: **Improving practice and understanding impact**: data-driven approaches and new data sources to create indicators for improving the effectiveness of peacebuilding practice that are useful for monitoring and evaluation of intervention such as real-time ICT applications, location-specific data, crowdsourced information, and closer interaction between humans and machines
	
	• **Room 903 - Locating hidden graves in Mexico using machine learning**
	  by Patrick Ball from Human Rights Data Analysis Group
	    Facilitator: Nigel Snoad, Verily Life Sciences
	
	• **Room 905 – Sentry**
	  by Dan Henebery from Hala Systems
	    Facilitator: Bas Bijlsma, Ministry of Foreign Affairs of Netherlands
	
	• **Room 907 – Real-time, public conflict data from ACLED**
	  by Roudabeh Kishi from Armed Conflict Location & Event Data Project (ACLED)
	    Facilitator: Christina Goodness, UN Department of Peacebuilding and Political Affairs-Department of Peace Operations
	
	• **Room 909 – Syria Mapping Project**
	  by Kate Keator from the Carter Center
	    Facilitator: Josje Spierings, Leiden University

16:30 – 17:30 Wrap-up & closing	Rapid-fire report back from the facilitator of each group and posting to the wall of ideas • Ayham Al Maleh Mapping suggestions for follow-up, including how best to connect going forward • Bas Bijilsma Closing of the seminar • Paige Arthur and Henk-Jan Brinkman