PEACE OPERATION DATA
Assessment, Analysis and Use

FINAL REPORT

Lead authors:
Evgenija Kroeker, University of Oxford
Maline Meiske, University of Oxford

Contributing authors:
Susanna Campbell, American University
Paige Arthur, Center on International Cooperation
Andrea Ruggeri, University of Oxford

Final report of the 2021/22 project ‘Peace Operation Data, Analytics, and Action’, a collaboration between American University, the Folke Bernadotte Academy, the University of Oxford, and the Center on International Cooperation
FOREWORD

Over the past two decades, the United Nations (UN) has increased its data collection commitment to monitor and evaluate the effect of UN peace operations. The UN has been gathering systematic information about the composition and location of civilian, police, and military personnel in peace operations, their activities, their locations, and conflict-related events involving these personnel. The Action for Peacekeeping (A4P) initiative, launched by the Secretary-General in 2018, and its 2021 implementation strategy, Action for Peacekeeping + (A4P+), have further spurred those efforts. The A4P+ Plan details specific results and deliverables for each of the A4P+ priorities, and a standalone Monitoring Framework presents quantitative and qualitative indicators to measure progress toward each priority, and the impact of UN peace operations more generally.

At the same time, academic researchers skilled in quantitative methods have made significant progress in analyzing the implementation, process, and impact of peace operations, using publicly available data. Although both the UN and academics want to evaluate the effectiveness of peace operations, these two communities have relatively little interaction.

The 2021/22 project ‘Peace Operation Data, Analytics, and Action’ – a collaboration between American University, the Folke Bernadotte Academy, the University of Oxford, and the Center on International Cooperation – aimed to fill this gap and bring together scholars and UN officials to identify best practices in current data-generation and analysis with important implications for UN peacekeeping and sustaining peace priorities. The project began with a workshop dedicated to providing an overview of the academic scholarship on peace operations and a discussion of challenges around the A4P+ Monitoring Framework. This was followed by three workshops on peacekeeping data and analytical methods: Third-Party Data; Community-Level Perception Surveys (presented by Dr. Patrick Vinck); and Artificial Intelligence, Natural Language Processing, and Text-as-Data (delivered by Dr. Margaret J. Foster). A final workshop reflected on the previous workshop discussions and examined the relationship between A4P, A4P+, decision-making, and peace operation effectiveness.

This final report of the ‘Peace Operation Data, Analytics, and Action’ project reflects and elaborates on the topics discussed during the workshops. It explores questions regarding the identification, collection, evaluation, and analysis of data, with the aim to improve how data is used and understood to monitor UN peace operations.
OVERVIEW OF THE REPORT

Section 1: From Concept to Measurements
Visualizes and exemplifies a stylized research design process. Starting with the idea of peacekeeping ‘impact’, it shows how to choose measures that reasonably link this concept with real-world indicators.

Section 2: Academic Data to Analyze Peacekeeping
Gives an overview of the state-of-the-art of third-party data compiled by quantitative peacekeeping scholars that can be used to study various characteristics of peace operations as well as their impact on a range of outcomes.

Section 3: Data Quality Challenges in Peacekeeping Data
Introduces the concept of ‘data quality’, its constitutive dimensions, explains how it might impact analytics, and provides guidance on how to evaluate the fitness for use of a particular dataset given the task at hand.

Section 4: Methods for Analyzing Peacekeeping Data
Gives an overview of and defines the steps in a typical data analytics process.

Appendix
- A) Comprehensive list of third-party datasets on UN peace operations
- B) Comprehensive list of third-party datasets on outcomes
- C) Slides from virtual workshop ‘Community-level Perception Surveys’, presented by Dr. Patrick Vinck on October 28, 2022
- D) Slides from virtual workshop ‘Artificial Intelligence (AI), Natural Language Processing (NLP), and Text-as-data’, presented by Dr. Margaret J. Foster on November 16, 2022
- E) References
1) FROM CONCEPT TO MEASUREMENTS

In this section, we explain and visualize a stylized research design process. Starting with the idea of peacekeeping ‘impact’, we describe how to choose measures that reasonably link this concept with real-world indicators.

➢ Step 1: From Background Concept to Systematized Concept

Before the impact of peace operations can be monitored and evaluated, it first needs to be clear what comprises ‘good’ peacekeeping performance. Are peacekeepers performing well when certain outcomes are achieved? This could, e.g., be the case when peacekeepers fulfil the mandates authorized by the UN Security Council and/or when violent conflict is mitigated and prevented. On the other hand, peacekeeping performance might be deemed ‘good’ when missions exhibit certain characteristics – for example, when peacekeepers are well-trained. Conceptualization – i.e. engaging with and choosing among the diverse meanings of ‘good’ performance – is a crucial first step that sets the scene for subsequent data-driven analytics. Without explicitly defining when peacekeepers are considered to be effective, it is challenging to develop qualitative or quantitative indicators and evaluate their suitability for the task at hand.

➢ Step 2: From Systematized Concept to Measurements

Once the dimensions of peace operation effectiveness have been selected and defined, observable implications can be developed through the selection of suitable measures which link the conceptual ideas about performance with real-world indicators. For example, when conflict mitigation and/or prevention is a dimension of ‘good’ performance, conflict intensity might be a reasonable indicator to evaluate peacekeeping effectiveness. On the other hand, when the fulfilment of women’s rights is an important dimension of impact, the degree of legal institutional protection for women might be a suitable measure. In general, indicators are only useful when they correlate with and represent the main ideas for the task at hand and are valid as well as reliable. Valid indicators measure what they are supposed to measure, whereas reliable indicators generate consistent results.

➢ Step 3: From Measurements to Scoring Cases

In the last step and depending on the type of indicator, scoring procedures are applied to the observations under study to produce values that can be meaningfully interpreted. As mentioned above, indicators for peacekeeping performance can be based on qualitative as well as quantitative information. For example, an indicator like conflict intensity might contain information on the number of battle fatalities over a certain time period. Then again, an indicator measuring civilians’ reliance on formal institutions in case of disputes might be based on more qualitative information. For instance, individual preferences extracted from on-the-ground perception surveys or via sentiment analysis of local social media posts.

The diagram below further visualizes this stylized research design process.
1) FROM CONCEPT TO MEASUREMENTS: A STYLIZED PROCESS

Background Concept

The broad constellation of meanings and understandings associated with a given concept.

- **Step 1: Conceptualization**
  Formulate a systematized concept with reference to the background concept.

- Background concepts are normally associated with a **variety of meanings**. A careful examination of these meanings helps clarify different understandings.
- When forming a systematized concept, a **choice** must be made among these options and the specific meaning defined.
- Emphasizing different aspects of a background concept is common. Any choice can be **justified** but must be linked to the goals and context of the task at hand.

- **Example**: Peacekeeping ‘impact’ / ‘effectiveness’
  - Certain outcomes are achieved
    - Mandates are implemented
    - Violent conflict is mitigated/prevented
    - Peaceful societies are built
  - Missions have certain attributes
    - Equipment is adequate
    - Peacekeepers are well-trained
    - Troop composition is diverse

Systematized Concept

A specific formulation of a background concept; commonly involves an explicit definition.

- **Step 2: Operationalization**
  Develop measures that link the systematized concept to real-world indicators.

- **Example**: “Peacekeeping is effective when it mitigates or prevents violent conflict”

- **Possible indicators**:²
  - Conflict intensity
    - Absence/low levels of violence
      - Battle-related deaths
      - Civilian casualties

---

¹ This visualization is taken from Adcock and Collier, 2001, and was adjusted for the purposes of this report.
² For a more thorough overview of this conceptual understanding in quantitative peacekeeping research see Kroeker and Ruggeri, 2022.
The suitability of an indicator is evaluated in relation to its validity and reliability.

Valid indicators measure what they are supposed to measure. They:

1. Capture the underlying systematized concept;
2. Produce values that correlate with similar measures of the same concept; and
3. Confirm common hypotheses when used.

Reliable indicators generate consistent values. They:

1. Produce the same values if scoring procedures are repeated over time;
2. Or by different people.

Invalid and unreliable indicators should not be used.

Measurements
Also called “indicators”. Includes any qualitative or quantitative scoring procedure that is systematic.

Step 3: Score cases
Apply selected scoring procedures to observations to produce meaningful values.

- Scoring procedures can be based on different types of data. Some measures encode purely mathematical values, whereas others encode more qualitative information.
- Qualitative, or categorical, indicators come in different forms. They represent one or more mutually exclusive categories within one indicator which can be meaningfully ordered.

Scores for cases
The values for cases generated by a particular measure. These include both numerical values and the results of qualitative classification.
2) ACADEMIC DATA TO ANALYSE PEACEKEEPING

Data-driven, quantitative studies on peace operations have been flourishing and growing in the last two decades. The expansion of experts and analyses has led to the creation of several new datasets and numerous empirical findings. Since the early 2000s, we know much more about the structure and organization of peace operations, and their impact on peace and security.

Many of the current data-generating efforts in academic research can be leveraged by UN initiatives to analyze and strengthen UN peacekeeping and sustaining peace priorities. In this section, we:

▪ Give an overview of the state of the art of third-party peacekeeping data used by quantitative scholars and their main findings;

▪ Highlight the usefulness of academic large-N data for UN initiatives that aim to strengthen peacekeeping as a tool for conflict management, including A4P and A4P+.

As an overview, the figure below shows the proportion of academic data suitable to study and evaluate each priority of the UN’s A4P and A4P+ framework. It is based on the 78 most widely used datasets in quantitative scholarship on peacekeeping. A detailed breakdown linking each dataset to A4P/A4P+ priorities can be found in the Appendix. We found that academic and other third-party data can be particularly valuable to assess progress on priorities relating to a collective political strategy, peacebuilding and sustaining peace, and the protection offered by peace operations.

*Figure 2.1 Coverage of A4P(+) Priorities in Academic Data*
A) ACADEMIC EFFORTS AT STUDYING PEACEKEEPING USING QUANTITATIVE DATA

WHAT ACADEMICS STUDY

- This figure depicts the dimensions of peacekeeping that have been studied with quantitative data. It is based on an analysis of the 120 most widely read and cited quantitative academic research articles published between 2000–2019.
- One third of the articles on peace operations focus on outcomes (e.g., violence reduction, conflict prevention, economic development, protection of civilians).
- Two thirds of the articles focus on the process or organization of peace operations (e.g., supply of peacekeepers, peacekeeper fatalities, UNSC mandates).

EXISTING ACADEMIC RESEARCH FINDS THAT

Peace/conflict outcomes:

- UN peacekeeping reduces the duration of conflict;
- Produces more durable peace;
- Geographically contains armed conflict;
- Protects civilians;
- Reduces battle-related violence;
- And lowers the odds of genocide over the long term.

More mixed findings for non-conflict outcomes:

- UN Peacekeeping seems to promote democratic processes;
- Stimulates economic demand during deployment, but economic growth rapidly declines when missions end;
- Only strengthens the rule of law during periods of peace, not while conflict is ongoing;
- Improves public health, e.g., maternal health outcomes and access to services;
- But no significant effect on other development indicators like literacy rate.

1 For an overview, see Di Salvatore and Ruggeri, 2017; Walter, Howard and Fortna, 2020.
3 Blair, Di Salvatore and Smidt, 2022; Beber et al., 2019; Blair, 2021; Gizelis and Cao, 2021; Kim 2017.
B) ACADEMIC DATA ON PEACE OPERATIONS

- As quantitative studies on peace operations have flourished, data collection efforts have followed suit. Many of these datasets are based on UN documents, e.g., personnel summaries, UNSC resolutions, or SG reports, but have been processed by either human coding or automated machine coding to allow for data-driven analysis.
- Early data collection efforts focused on the presence of peace operations and personnel numbers.
- More recently, we have seen an extension to other facets of peacekeeping and more fine-grained data. New facets include leadership features, tasks assigned to missions, violence against peacekeepers and deployment speed. Finer-grained data arrived at the subnational level, gender-disaggregated, at the monthly level and even studying events instead of larger temporal spells.
- From a mere study of the presence of peacekeepers, we now know much more about the effects of size, composition, and actions of peacekeepers.

Table 2.1 Sample of Datasets on UN Peace Operations

<table>
<thead>
<tr>
<th>Topic</th>
<th>Dataset</th>
<th>Content</th>
<th>Scope</th>
<th>Unit of Analysis</th>
<th>Sources</th>
<th>Regular Updates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supply of Peace-keepers</td>
<td>International Peace Institute (IPI) Data</td>
<td>Uniformed personnel contributions (troop, police, observer)</td>
<td>UN POs; 1990–2018</td>
<td>Diverse, i.e., country-month, mission-month</td>
<td>UN's monthly mission personnel summaries</td>
<td>yes</td>
</tr>
<tr>
<td>Mandates</td>
<td>Peace-keeping Mandates (PEMA) Dataset</td>
<td>Initial and revised PO mandates, covering 39 different tasks</td>
<td>UN POs (Africa); 1991–2017</td>
<td>PO-UNSC resolution</td>
<td>UNSC resolutions</td>
<td>no</td>
</tr>
<tr>
<td>Subnational Deployment</td>
<td>Geocoded Peace-keeping Operations (Geo-PKO) Dataset</td>
<td>PO deployment at local level, size, troop type, headquarters, T/PCCs</td>
<td>UN POs (Africa); 1994–2020</td>
<td>GIS coordinates-month</td>
<td>Mission deployment maps; UNSG mission progress reports</td>
<td>yes</td>
</tr>
<tr>
<td>Safety of Peace-keepers</td>
<td>UCDP Peace-makers at Risk (PAR) Dataset</td>
<td>Geo-referenced events on violence against peacekeepers</td>
<td>UN &amp; non-UN POs (Africa); 1989–2009</td>
<td>Event</td>
<td>Factiva (News archive database); UN, NGO and open-source reports</td>
<td>no</td>
</tr>
</tbody>
</table>
C) OUTCOME & PEACEKEEPING PERFORMANCE DATA

- Whether peacekeeping works has become one of the most important questions studied by quantitative researchers in the past decade. Predominantly based on analyses of easily accessible and comprehensive ‘off-the-shelf’ datasets, peacekeeping performance has been evaluated against a wide range of outcomes.
- Earlier studies focused on the ability of peacekeepers to manage conflict, that is to curb violence, increase the duration of post-conflict peace or prevent the recurrence of conflict.
- More recent work focuses on the impact of peacekeeping on more sustainable forms of peace, including democratization processes, the rule of law and economic development.
- Challenges remain regarding the availability of data on indicators of positive peace and fine-grained data, as well as a clear understanding about the interdependency of mission tasks.

Table 2.2 Sample of Datasets on Outcomes

<table>
<thead>
<tr>
<th>Topic</th>
<th>Dataset</th>
<th>Content</th>
<th>Scope</th>
<th>Unit of Analysis</th>
<th>Sources</th>
<th>Regular Updates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conflict</td>
<td>Uppsala Conflict Data Program (UCDP)</td>
<td>Various datasets on conflict, incl. armed conflict; conflict termination; non-state actor; one-sided violence; georeferenced events; dyadic data</td>
<td>Global; 1945/1989–2020</td>
<td>Conflict-group-year; GIS coordinates-month</td>
<td>Several, incl. global news, local news, NGO reports</td>
<td>yes</td>
</tr>
<tr>
<td>Political Violence</td>
<td>Social Conflict Analysis Database (SCAD)</td>
<td>Forms of social conflict, incl. protests, riots, strikes, inter-communal conflict, government violence against civilians</td>
<td>Africa, Latin American and the Caribbean; 1990–2017</td>
<td>Date-locality</td>
<td>Lexis-Nexis</td>
<td>yes</td>
</tr>
<tr>
<td>Terrorism</td>
<td>Global Terrorism Database (GTD)</td>
<td>Information on domestic and international terrorist attacks</td>
<td>Global; 1970–2019</td>
<td>GIS coordinates-month</td>
<td>News</td>
<td>yes</td>
</tr>
<tr>
<td>Democracy</td>
<td>Varieties of Democracy (V-Dem)</td>
<td>Indicators on electoral, liberal, participatory, deliberative and egalitarian democracy</td>
<td>Global; 1789–2020</td>
<td>Country-year</td>
<td>Third-party data; country experts</td>
<td>yes</td>
</tr>
</tbody>
</table>
D) TIPS WHEN USING THIRD-PARTY DATA

Selecting a suitable dataset to score cases depends on the objectives of the study, the research question, and the overall requirements of the project. When choosing from existent academic and other third-party data, several factors need to be taken into consideration regarding the suitability and limitations of the data:

- **Relevance**: Relevance is concerned with the degree to which the data meet the needs of the data user. It looks at whether data are applicable to measure the concept(s) being studied, and whether they are sufficiently complete, consistent, and uniform to fulfill the task at hand.
  
  ? Are the data suitable for the intended task?

- **Scope**: Scope looks at the amount of data or potential reach available. Datasets can vary, particularly with regard to their temporal and geographic scope. If a dataset is to be used over a longer period of time, a related factor to consider is whether the data is regularly updated and, if so, in which intervals.
  
  ? Is the available temporal and geographic scope and update frequency sufficient for the purpose of the analysis?

- **Aggregation**: Aggregation refers to the analytical level at which data are collected or presented – both in geographical (e.g., national vs local) and temporal (e.g., year vs month) terms. Data aggregation can be problematic as it can obscure individual effects of interventions or policies.
  
  ? Is the level of data aggregation – geographic area and time – appropriate for the assessment that needs to be made?

- **Source**: A data source is the place or person from where the data that is being used originates. Datasets might source their data from other publicly available datasets, population surveys, expert surveys, or news reports. Third-party data providers are not necessarily collecting the data on their own; and sources might vary from variable to variable.
  
  ? Where does the data come from; and is the source reliable?

- **Usability**: Data usability considers the extent that all necessary data are easily accessible, that definitions are clear and the data is easily interpretable; the ability for a user to derive useful information; and the extent to which data can contribute to actionable analytics while reducing the risk of misinterpretation.
  
  ? Is the data easily accessible, understandable, and transformable for actionable use?

- **Quality**: Data quality assesses the conditions of the data based on factors such as accuracy, reliability, completeness, comparability, and timeliness. Measuring data quality levels can help to identify data errors and biases and assess whether the data serve their intended purpose.
  
  ? Are the data of a high enough quality?
3) DATA QUALITY CHALLENGES IN PEACEKEEPING DATA

Data quality is a critical issue for data-driven analysis. Accessible, complete, and accurate data are a key component for sound analysis, whereas poor-quality input data produce faulty and misleading research and can lead to poor policy recommendations.

Data quality describes a certain condition of a given dataset, its ‘fitness for use’ or, put differently, the ‘degree to which dimensions of data meet certain requirements.’ Determining whether a dataset is of high quality involves examining its characteristics and deciding whether it meets the needs of the task at hand. In the following, this section introduces key dimensions of data quality;¹ it explores one example of data quality challenges – completeness – in more detail; and it provides guiding questions and common data quality checks.

A) DIMENSIONS OF DATA QUALITY

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>The extent to which data are correct and reliable. Or if the data values stored in the database correspond to real-world values.</td>
</tr>
<tr>
<td>Completeness</td>
<td>The degree to which values are included for all cases in a data collection effort (i.e., not missing). Or the extent to which data are of sufficient breadth, depth, and scope for the task at hand.</td>
</tr>
<tr>
<td>Comparability</td>
<td>The degree to which data values have the same definition and are measured in the same way across cases.</td>
</tr>
<tr>
<td>Consistency</td>
<td>The degree to which data values of a set of attributes comply with a rule.</td>
</tr>
<tr>
<td>Timeliness</td>
<td>The degree to which data is up to date.</td>
</tr>
<tr>
<td>Uniqueness</td>
<td>Uniqueness is a measure of the number of duplicates; data is unique if it appears only once in a data set.</td>
</tr>
<tr>
<td>Validity</td>
<td>Validity is concerned with whether an indicator measures what it is supposed to measure; measurement validity concerns the link between empirical reality and concepts.</td>
</tr>
</tbody>
</table>

¹ For further dimensions of data quality, see: Black and van Nederpelt, 2020; Pipino, Lee and Yang, 2002; Wand and Wang, 1996; Wang and Strong, 1996.
Lack of data availability and missing data are common challenges in the social sciences.

Take the World Bank’s World Development Indicators (WDI) as an example. The WDI are the most extensive, current, and reliable source for socio-economic and development data. They contain over 1,443 indicators for 266 countries and regional and global aggregations over the 1960–2020 period (as of March 2022). However, 66% of its possible observations are missing, which drops to 53% for the post-Cold War period.

Figure 3.1 maps the average percent of data completeness for each country (over all indicators, between 1989–2020). The highest level of data completeness is found in South America with around 55%, the lowest level in Africa with 40%, and the Western countries lie in between.

Missing observations do not only limit the available sample and data for the analysis, but, even more importantly, can also entail issues of non-random missing data. This means that there is a systematic relationship between the propensity of missing values and other observed variables, or even between the propensity of a value to be missing and its values.

Understanding why observations are missing is important for handling the remaining data correctly. A sample may still be representative of the population if data are missing completely at random. But non-random, systematically missing data may lead to biased analysis. Thus, it is important to check whether data are missing due to some specific events, features of location, or other possible systematic biases.

Several options exist for dealing with missing data. Basic approaches include analysis of the complete cases and associated weighting methods, and methods that impute the missing values. More principled methods are based on statistical models and likelihood-based approaches.

---

1 This figure is drawn from Meiske and Ruggeri, 2022.
C) GUIDELINES FOR ASSESSING DATA QUALITY

GUIDING QUESTIONS FOR SELECTING AND INSPECTING DATA INCLUDE:

▪ Are the data produced over time and across cases measured **consistently**?
▪ Can we **compare** the data collected over time and in different peace operations?
▪ Are our data complete or are there **missing observations**?
▪ Are the data missing due to some specific events, features of location, or other possible **systematic biases**?
▪ What is the degree of **measurement error** in our data?
▪ Are the phenomena that we want to assess for peace operations’ activities and outcomes measured with **apt proxies**?

COMMON DATA QUALITY CHECKS INCLUDE:

▪ Check that each data record only occurs once in the dataset used and identify **potential duplicates**.
▪ Ascertain the primary data sources, data collection method(s), and definitions to determine whether data are **comparable over time and place**.
▪ Check how recent the latest data entries are or when they were **last updated** to determine timeliness.
▪ Apply **formatting checks**, e.g., for mismatched data types or variations in how values are entered, to ensure consistency.
▪ Check for **missing values** and **null values** to identify and fix issues of data completeness.
▪ In case of missingness, assess whether data are **missing at random** or not by exploring correlations with related variables and trends between groups and over time.
▪ Examine and/or **plot the data** distribution for each variable to assess integrity.
▪ Look for **outliers**, or extreme values.
▪ Evaluate how closely the **data correlate** with other, related measures to test for accuracy and validity.
▪ **Take a random sample** of the data and **visually inspect** it.
4) METHODS FOR ANALYSING PEACEKEEPING DATA

Information related to peace operations can be leveraged via data analytics. Data analytics is the process of extracting relevant and actionable information from data in view of a particular decision-making problem. Analyzing data is especially helpful to:

▪ Understand what happened (description);
▪ Why it happened (diagnosis);
▪ What might happen in the future (prediction); and
▪ What should be done about it (recommendation).

STYLIZED DATA ANALYTICS PROCESS

1 Define what type of data is needed, keeping in mind what measures have been chosen to approximate the systematized concept for the task at hand.

2 Collect the necessary data. Multiple approaches exist:
   - Hand-coding: Manually collect information from online as well as offline sources
   - Natural language processing: Use algorithms that understand human language and turn text into data
   - Surveys: Administer new information for task at hand via direct participation of a specified population

3 Inspect, clean, and prepare the data collected for subsequent analysis. Make sure to ascertain the quality of the data collected (see Section 3).

4 Analyze and interpret the data, using different techniques:
   (1) Description
       • Visualize key indicators (charts, graphs, maps, dashboards)
   (2) Diagnosis
       • Explore relation between indicators via statistical analysis
   (3) Prediction
       • Simulate scenarios based on historic patterns and current events
   (4) Recommendation
       • Based on results of preceding types of data analytics, what ought to be the best course of action?

1 See Appendix D for further information on how these text-as-data techniques work.
2 See Appendix C for further information on how to administer community-level perception surveys.
# A) COMPREHENSIVE LIST OF THIRD-PARTY DATASETS ON UN PEACE OPERATIONS

<table>
<thead>
<tr>
<th>A4P(+) Priority</th>
<th>Topic</th>
<th>Dataset</th>
<th>Link</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A4P 3</strong>: protection;</td>
<td>(Sub-national) Deployment</td>
<td>Geocoded Peacekeeping Operations (Geo-PKO) Dataset</td>
<td><a href="https://www.pcr.uu.se/data/geo-pko/">Link</a></td>
</tr>
<tr>
<td><strong>A4P+ 2</strong>: strategic &amp; operational integration;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>A4P 7</strong>: cooperation with host countries</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>A4P 3</strong>: protection;</td>
<td>(Sub-national) Deployment</td>
<td>Robust African Deployment of Peacekeeping Operations (RADPKO)</td>
<td><a href="https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/BQU5VD">Link</a></td>
</tr>
<tr>
<td><strong>A4P+ 2</strong>: strategic &amp; operational integration;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>A4P+ 7</strong>: cooperation with host countries</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>A4P+ 3</strong>: capabilities &amp; mindset</td>
<td>(Sub-national) Deployment</td>
<td>Peace Operation Deployment Speed</td>
<td><a href="https://www.prio.org/journals/jpr/replicationdata">Link</a></td>
</tr>
<tr>
<td><strong>A4P 7</strong>: partnerships;</td>
<td>(Sub-national) Deployment</td>
<td>Third-Party Peacekeeping in Intrastate Disputes</td>
<td><a href="https://uca.edu/politicalscience/dadm-project/dadm-data-sets/">Link</a></td>
</tr>
<tr>
<td><strong>A4P+ 1</strong>: coherent political strategy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>A4P 7</strong>: partnerships;</td>
<td>(Sub-national) Deployment</td>
<td>Military and Non-Military Interventions Dataset (MILINDA)</td>
<td><a href="https://lehrstuhlib.uni-goettingen.de/sonstiges/milinda/">Link</a></td>
</tr>
<tr>
<td><strong>A4P+ 1</strong>: coherent political strategy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>A4P 7</strong>: partnerships;</td>
<td>(Sub-national) Deployment</td>
<td>International Military Intervention (IMI) Data</td>
<td><a href="https://www.k-state.edu/polsci/intervention/">Link</a></td>
</tr>
<tr>
<td><strong>A4P+ 1</strong>: coherent political strategy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>A4P 1</strong>: politics;</td>
<td>Mandates</td>
<td>Tasks Assigned to Missions in their Mandates (TAMM)</td>
<td><a href="http://www.gabriellloyd.com/tamm">Link</a></td>
</tr>
<tr>
<td><strong>A4P+ 6</strong>: strategic communications</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>A4P 1</strong>: politics;</td>
<td>Mandates</td>
<td>Peacekeeping Mandates (PEMA) Dataset</td>
<td>[available soon]</td>
</tr>
<tr>
<td><strong>A4P+ 6</strong>: strategic communications</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------</td>
<td>---------</td>
<td>---------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>A4P 1: politics; A4P 7: partnerships; A4P+ 1: coherent political strategy; A4P+ 6: strategic communications</td>
<td>Mandates</td>
<td>Peace Operation Mandates (POM) Data</td>
<td>[available soon]</td>
</tr>
<tr>
<td>A4P 3: protection; A4P 5: performance and accountability</td>
<td>Supply of Peacekeepers</td>
<td>UN PO Personnel Shortfall Data</td>
<td><a href="https://www.prio.org/journals/jpr/replicationdata">https://www.prio.org/journals/jpr/replicationdata</a></td>
</tr>
<tr>
<td>A4P 7: partnerships; A4P+ 1: coherent political strategy</td>
<td>Supply of Peacekeepers</td>
<td>SIPRI Multilateral Peace Operations Database</td>
<td><a href="https://www.sipri.org/databases/pko">https://www.sipri.org/databases/pko</a></td>
</tr>
<tr>
<td>A4P 7: partnerships; A4P+ 1: coherent political strategy</td>
<td>Supply of Peacekeepers</td>
<td>UN &amp; non-UN Peacekeeping Dataset</td>
<td><a href="https://www.corinnebara.net/data/">https://www.corinnebara.net/data/</a></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>A4P 4: safety &amp; security; A4P+ 4: accountability to peacekeepers</td>
<td>Safety of Peacekeepers</td>
<td>UCDP Peacemakers at Risk (PAR) Dataset</td>
<td><a href="https://www.prio.org/journals/jpr/replicationdata">https://www.prio.org/journals/jpr/replicationdata</a></td>
</tr>
<tr>
<td>A4P 8: conduct of peacekeepers and peacekeeping operations; A4P+ 5: accountability of peacekeepers</td>
<td>SEA by Peacekeepers</td>
<td>Sexual Exploitation and Abuse in Peacekeeping Missions Data</td>
<td><a href="https://www.prio.org/journals/jpr/replicationdata">https://www.prio.org/journals/jpr/replicationdata</a></td>
</tr>
<tr>
<td>A4P 8: conduct of peacekeepers and peacekeeping operations; A4P+ 5: accountability of peacekeepers</td>
<td>SEA by Peacekeepers</td>
<td>Sexual Exploitation and Abuse by Peacekeepers (SEAP) Dataset</td>
<td><a href="https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/CRQLIU">https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/CRQLIU</a></td>
</tr>
<tr>
<td>--------------------------------------------------------</td>
<td>-----------------</td>
<td>-----------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
## B) COMPREHENSIVE LIST OF THIRD-PARTY DATASETS ON OUTCOMES

<table>
<thead>
<tr>
<th>A4P(+) Priority</th>
<th>Topic</th>
<th>Dataset</th>
<th>Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>A4P 1: politics; A4P 3: protection; A4P 6: peacebuilding and sustaining peace</td>
<td>Conflict</td>
<td>Uppsala Conflict Data Program (UCDP) datasets</td>
<td><a href="https://ucdp.uu.se/downloads/">https://ucdp.uu.se/downloads/</a></td>
</tr>
<tr>
<td>A4P 1: politics; A4P 7: partnerships; A4P+ 1: collective coherence behind a political strategy</td>
<td>Conflict, diplomacy</td>
<td>Diplomatic Intervention and Civil War data</td>
<td><a href="https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/24822">https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/24822</a></td>
</tr>
<tr>
<td>A4P 6: peacebuilding and sustaining peace</td>
<td>Democracy, political rights</td>
<td>Freedom House Index</td>
<td><a href="https://freedomhouse.org/reports/publication-archives">https://freedomhouse.org/reports/publication-archives</a></td>
</tr>
<tr>
<td>A4P 2: women, peace and security; A4P 6: peacebuilding and sustaining peace</td>
<td>Development</td>
<td>Clio Infra</td>
<td><a href="https://clio-infra.eu">https://clio-infra.eu</a></td>
</tr>
<tr>
<td>A4P 6: peacebuilding and sustaining peace</td>
<td>Development</td>
<td>Gapminder</td>
<td><a href="https://www.gapminder.org/data/">https://www.gapminder.org/data/</a></td>
</tr>
<tr>
<td>A4P 1: politics; A4P 3: protection; A4P 6: peacebuilding and sustaining peace</td>
<td>Displaced persons</td>
<td>Forcibly Displaced Populations</td>
<td><a href="https://www.systemicpeace.org/inscrdata.htm">https://www.systemicpeace.org/inscrdata.htm</a> l</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>A4P 1: politics; A4P 3: protection; A4P 6: peacebuilding and sustaining peace</td>
<td>Displaced persons</td>
<td>The Refugee Project</td>
<td><a href="https://therefugeeproject.org/#/2020">https://therefugeeproject.org/#/2020</a></td>
</tr>
<tr>
<td>A4P 1: politics; A4P 3: protection; A4P 6: peacebuilding and sustaining peace</td>
<td>Displaced persons</td>
<td>Internal Displacement Monitoring Centre (IDMC)</td>
<td><a href="https://data.humdata.org/organization/international-displacement-monitoring-centre-idmc">https://data.humdata.org/organization/international-displacement-monitoring-centre-idmc</a></td>
</tr>
<tr>
<td>A4P 6: peacebuilding and sustaining peace</td>
<td>Economy</td>
<td>Penn World Table</td>
<td><a href="https://www.rug.nl/ggd/c/productivity/pwt/?lan">https://www.rug.nl/ggd/c/productivity/pwt/?lan</a> g=en</td>
</tr>
<tr>
<td>A4P+ 4: accountability to peacekeepers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A4P+ 4: accountability to peacekeepers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A4P 6: peacebuilding and sustaining peace</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A4P 2: women, peace and security;</td>
<td>Health</td>
<td>Demographic and Health Surveys (DHS)</td>
<td><a href="https://dhsprogram.com/data/">https://dhsprogram.com/data/</a></td>
</tr>
<tr>
<td>A4P 6: peacebuilding and sustaining peace</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A4P 6: peacebuilding and sustaining peace</td>
<td>Human rights</td>
<td>Cingranelli-Richards (CIRI) Human Rights Data Set</td>
<td><a href="https://dataverse.harvard.edu/dataverse/cirihumanrightsdata">https://dataverse.harvard.edu/dataverse/cirihumanrightsdata</a></td>
</tr>
<tr>
<td>A4P 1: politics;</td>
<td>Human rights</td>
<td>Political Terror Scale (PTS)</td>
<td><a href="https://www.politicalterrorism.org/Data/Download.html">https://www.politicalterrorism.org/Data/Download.html</a></td>
</tr>
<tr>
<td>A4P 6: peacebuilding and sustaining peace</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A4P 6: peacebuilding and sustaining peace</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A4P 1: politics;</td>
<td>Military</td>
<td>Stockholm International Peace Research Institute (SIPRI) databases</td>
<td><a href="https://www.sipri.org/databases">https://www.sipri.org/databases</a></td>
</tr>
<tr>
<td>A4P 3: protection;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A4P+ 7: cooperation with host countries</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A4P 1: politics;</td>
<td>Political events</td>
<td>GDELT Event Database</td>
<td><a href="https://www.gdeltproject.org/data.html">https://www.gdeltproject.org/data.html</a></td>
</tr>
<tr>
<td>A4P 6: peacebuilding and sustaining peace</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A4P Code</td>
<td>Description</td>
<td>Category</td>
<td>Source/Link</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
<td>----------</td>
<td>-------------</td>
</tr>
<tr>
<td>A4P 1</td>
<td>politics</td>
<td>Political events</td>
<td>Keesing's Record of World Events <a href="https://library.princeton.edu/resource/3894">https://library.princeton.edu/resource/3894</a></td>
</tr>
<tr>
<td>A4P 1</td>
<td>politics</td>
<td>Political violence</td>
<td>Major Episodes of Political Violence <a href="https://www.systemicpeace.org/inscrdata.htm">https://www.systemicpeace.org/inscrdata.htm</a></td>
</tr>
<tr>
<td>A4P 1</td>
<td>politics</td>
<td>Political violence</td>
<td>Social Conflict Analysis Database (SCAD) <a href="https://www.strausscenter.org/ccaps-research-areas/social-conflict/database/">https://www.strausscenter.org/ccaps-research-areas/social-conflict/database/</a></td>
</tr>
<tr>
<td>A4P 6</td>
<td>peacebuilding and sustaining peace</td>
<td>Public opinion</td>
<td>Afrobarometer <a href="https://afrobarometer.org/data/merged-data">https://afrobarometer.org/data/merged-data</a></td>
</tr>
<tr>
<td>A4P 6</td>
<td>peacebuilding and sustaining peace</td>
<td>Public opinion</td>
<td>AmericasBarometer (LAPOP) <a href="https://www.vanderbilt.edu/lapop/raw-data.php">https://www.vanderbilt.edu/lapop/raw-data.php</a></td>
</tr>
<tr>
<td>A4P 6</td>
<td>peacebuilding and sustaining peace</td>
<td>Public opinion</td>
<td>Eurobarometer <a href="https://europa.eu/eurobarometer/surveys/browse/all">https://europa.eu/eurobarometer/surveys/browse/all</a></td>
</tr>
<tr>
<td>A4P 6</td>
<td>peacebuilding and sustaining peace</td>
<td>Public opinion</td>
<td>Asian Barometer <a href="http://www.asianbarometer.org/data">http://www.asianbarometer.org/data</a></td>
</tr>
<tr>
<td>A4P 4</td>
<td>safety and security</td>
<td>State authority</td>
<td>Political Instability Task Force (PITF) State Failure Problem Set <a href="https://www.systemicpeace.org/inscrdata.htm">https://www.systemicpeace.org/inscrdata.htm</a></td>
</tr>
<tr>
<td><strong>A4P 1:</strong> politics</td>
<td>State authority</td>
<td>Coups d'Etat</td>
<td><a href="https://www.systemicpeace.org/inscrdata.htm">https://www.systemicpeace.org/inscrdata.htm</a>!</td>
</tr>
<tr>
<td>---------------------</td>
<td>----------------</td>
<td>--------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td><strong>A4P 1:</strong> politics; <strong>A4P 4:</strong> safety and security; <strong>A4P+ 4:</strong> accountability to peacekeepers</td>
<td>State authority</td>
<td>State Fragility Index and Matrix, Time-Series Data</td>
<td><a href="https://www.systemicpeace.org/inscrdata.htm">https://www.systemicpeace.org/inscrdata.htm</a>!</td>
</tr>
<tr>
<td><strong>A4P 6:</strong> peacebuilding and sustaining peace</td>
<td>State authority</td>
<td>Counterbalancing / Coup-Proofing Data</td>
<td><a href="https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/JCVR1H">https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/JCVR1H</a></td>
</tr>
<tr>
<td><strong>A4P 1:</strong> politics</td>
<td>Subnational data</td>
<td>PRIO-GRID data</td>
<td><a href="https://grid.prio.org/#/download">https://grid.prio.org/#/download</a></td>
</tr>
<tr>
<td><strong>A4P 1:</strong> politics</td>
<td>Terrorism</td>
<td>High Casualty Terrorist Bombings</td>
<td><a href="https://www.systemicpeace.org/inscrdata.htm">https://www.systemicpeace.org/inscrdata.htm</a>!</td>
</tr>
<tr>
<td><strong>A4P 1:</strong> politics</td>
<td>Terrorism</td>
<td>Global Terrorism Database (GTD)</td>
<td><a href="https://www.start.umd.edu/gtd/access/">https://www.start.umd.edu/gtd/access/</a></td>
</tr>
<tr>
<td><strong>A4P 1:</strong> politics; <strong>A4P 6:</strong> peacebuilding and sustaining peace</td>
<td>Violence</td>
<td>Global Violent Deaths (GVD)</td>
<td><a href="https://www.smallarmssurvey.org/database/global-violent-deaths-gvd">https://www.smallarmssurvey.org/database/global-violent-deaths-gvd</a></td>
</tr>
<tr>
<td><strong>A4P 1:</strong> politics; <strong>A4P 6:</strong> peacebuilding and sustaining peace</td>
<td>Violence</td>
<td>Targeted Mass Killing (TMK) dataset</td>
<td><a href="https://politicsir.cass.anu.edu.au/about-targeted-mass-killing-dataset">https://politicsir.cass.anu.edu.au/about-targeted-mass-killing-dataset</a></td>
</tr>
<tr>
<td><strong>A4P 1:</strong> politics; <strong>A4P 6:</strong> peacebuilding and sustaining peace</td>
<td>War, military, diplomacy, alliances</td>
<td>Correlates of War (COW) datasets</td>
<td><a href="https://correlatesofwar.org/data-sets">https://correlatesofwar.org/data-sets</a></td>
</tr>
</tbody>
</table>
C) SLIDES FROM VIRTUAL WORKSHOP ‘COMMUNITY-LEVEL PERCEPTION SURVEYS’, PRESENTED BY DR. PATRICK VINCK ON OCTOBER 28, 2022

VIRTUAL PEACEKEEPING DATA WORKSHOP #4

COMMUNITY-LEVEL PERCEPTION SURVEYS

BRIDGING SCHOLARSHIP AND PRACTICE

PATRICK VINCK

PeacebuildingData.org
"Why can’t those people leave peace to the experts?"

Making available detailed and reliable up to date data
Produced by quality teams from local universities

In order to
Monitor and assess key indicators on peacebuilding and progress linked to ongoing interventions

Which leads to...
Understand the situation and make good decisions to break the cycle of violence and build peace
COMMON DATA LANDSCAPE
GAPS
• LOCAL
• GLOBAL

WHAT IS HAPPENING HERE?

COMMON DATA LANDSCAPE
GAPS
• TEMPORAL

WHAT HAPPENED HERE?
WHAT IS HAPPENING HERE?
“Get me everything on everybody.” © NEWYORKER
**DRC example**

- **25+ QUARTERLY POLLS**
- **25 SAMPLING STRATA**
- **>125,000 INTERVIEWS 2014-2022**

**UNIQUE LONGITUDINAL DATASET**

**RIGOROUS METHODOLOGY IN COMPLEX SETTING**
- MIXED METHODS – TECHNOLOGY – LONGITUDINAL

**CREATING EVIDENCE-BASED POLICY SPACE**
- PARTICIPATORY APPROACH – ENGAGEMENT – RECOMMENDATIONS

---

**LESSONS LEARNED**

- "CREATE" DATA
  - SAMPLING
  - REDUCE "EXTRACTIVE" RESEARCH PRACTICES
  - REDUCE COST OF SURVEYS
  - EXAMPLES

- CREATE SPACE FOR POLICY DISCUSSIONS INFORMED BY DATA, EVIDENCE

---

*PeaceBuildingData.org*
WHAT ARE WE MEASURING?

- **Trust in MONUSCO to ensure security (% yes)**
  - Total: 35% yes, 5% no
- **MONUSCO’s contribution to your security (% none)**
  - Total: 29% poor, 73% average

_DATA FROM JULY – AUGUST 2019_
Figure 19: Population trust in FARDC to ensure security (% yes)

Sentiment de Sécurité
% en sécurité – très en sécurité

Activités quotidiennes/ par province

Tendances provinciales (Total pour NK, SK, IT)
**Acteurs de sécurité**

% confiance pour assurer la sécurité

Dernières données: mars 2021

Tendances provinciales (Total pour NK, SK, IT)

**RELATIONS ETHNIQUES**

% bonne – très bonne

Dernières données: mars 2021

Tendances provinciales (Total pour NK, SK, IT)
SSU / Stabilization indicators

- Political participation
- Inclusion / representation by local authorities
- Trust in authorities, security actors, within community
- Satisfaction with services
- ...

### Perception of humanitarians resp. ebola

Perception of humanitarian actors responding to Ebola (% positive)

<table>
<thead>
<tr>
<th></th>
<th>Goma</th>
<th>Beni</th>
<th>BumbeBo</th>
<th>Katwa</th>
<th>Bukavu</th>
<th>Uvira</th>
<th>Bunia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability to assist those who need it most</td>
<td>24%</td>
<td>51%</td>
<td>10%</td>
<td>11%</td>
<td>5%</td>
<td>4%</td>
<td>24%</td>
</tr>
<tr>
<td>Ability to respond rapidly</td>
<td>30%</td>
<td>55%</td>
<td>18%</td>
<td>13%</td>
<td>5%</td>
<td>2%</td>
<td>27%</td>
</tr>
<tr>
<td>Behavior toward the community</td>
<td>35%</td>
<td>51%</td>
<td>9%</td>
<td>8%</td>
<td>10%</td>
<td>1%</td>
<td>23%</td>
</tr>
<tr>
<td>Respect customs and traditions</td>
<td>22%</td>
<td>39%</td>
<td>2%</td>
<td>0%</td>
<td>4%</td>
<td>0%</td>
<td>11%</td>
</tr>
<tr>
<td>How they involve local authorities</td>
<td>26%</td>
<td>50%</td>
<td>19%</td>
<td>11%</td>
<td>3%</td>
<td>0%</td>
<td>23%</td>
</tr>
<tr>
<td>Ability to influence decisions of humanitarians</td>
<td>8%</td>
<td>12%</td>
<td>1%</td>
<td>1%</td>
<td>5%</td>
<td>0%</td>
<td>4%</td>
</tr>
</tbody>
</table>
LESSONS LEARNED

WHO?

25 territoires / cities (strata)

In each strata, random selection of 9 groupements (quartiers)

In each groupement, random selection of 3 villages (avenues)

In each village random selection of 8 households (geographic EPI method)

In each household, random selection of 1 adult
### I4S Priority Zones

<table>
<thead>
<tr>
<th>Zone Prioritaire</th>
<th>Sous Zone</th>
<th>Chefferies/ Secteurs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Sud Irumu</td>
<td>1.1 Sud</td>
<td>Bahema Mbinga</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bahema Bega</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kamasi Tshabi</td>
</tr>
<tr>
<td>1.2 Sud Est</td>
<td></td>
<td>Bahema Sud</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wacondu Bindi</td>
</tr>
<tr>
<td>2 Plaine Ruzizi</td>
<td>2.1 Uvira</td>
<td>Plaine de Ruzi</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bavira</td>
</tr>
<tr>
<td>2.2 Mwiriga</td>
<td></td>
<td>Secteur d’Homeware</td>
</tr>
<tr>
<td>2.3 Uvira Ville</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Kibanganga</td>
<td>3.1 Ruzhuni</td>
<td>Sseko</td>
</tr>
<tr>
<td></td>
<td>3.2 Masi-Rings</td>
<td>Bashelli</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ave Rings Ruzhuni**</td>
</tr>
<tr>
<td>4 Kalaha-Buhavu</td>
<td>4.1 Hadi Plateau</td>
<td>Buhavu**</td>
</tr>
<tr>
<td></td>
<td>4.2 Ulototo</td>
<td>Buhavu**</td>
</tr>
<tr>
<td>5 Mambasa</td>
<td>5.1 Sud Ouest</td>
<td>Babila Balumbi</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Babila Balubisa</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Baribika</td>
</tr>
<tr>
<td></td>
<td>5.2 Nord Est</td>
<td>Babila Balawina</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mambasa</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Walone Kiro</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Walone Dose</td>
</tr>
<tr>
<td>6 Beri</td>
<td>6.1 Nord</td>
<td>Beri Mboi</td>
</tr>
<tr>
<td>6.2 Est</td>
<td></td>
<td>Kwimpangi</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Watalinga</td>
</tr>
</tbody>
</table>

**'MONUSCO STUDY'**

Comprehensive Evaluation and Proximity-based Sentiment Analysis
LESSONS LEARNED
HOW?

MIXED METHODS

PRACTICAL, IN-DEPTH, BUT ALSO QUANTIFIABLE UNDERSTANDING OF SOCIAL ISSUES WITH DIRECT PRACTICAL OUTCOMES SHAPING PROGRAMS AND POLICIES.

Figure 2. Mixing these methods should be safe.
Illustration by Patrick Heck.
Overall strategy?

- Crowdsourcing?
- Crowdseeding?
- Crowdseeding neighborhood method
- Phone-based interviews
- Face to face (multiple sources)

- Reliable
- Timely
- Representative
- Trends
- Participatory

---

American University | Folke Bernadotte Academy | University of Oxford | Center on International Cooperation
WHAT TO ASK?

Domains of Measurement
(Core & Country Specific)

Dimensions

“Levels”
Methods

Concept Maps

Drivers
- Structural issues (access to land, services)
- Nature
- Education
- Violence
- Conflict in schools | in community
- Access to education

Interdependence
- Participation
- Change

Social Cohesion
- Social justice
- Social conflict
- Social relations

Security
- (physical, economic, political)

Trust
- Equity
- Individual Rights
- Tolerance
- Friends / Family
- Community
- State Institutions

Civil Engagement
- Participation

Methods - Personas

Personas

Person of interest - Name:

Location
- Gender
- Age
- Occupation
- Economic Level

Income class / Wealth
- Other characteristics:

Risk and Vulnerabilities:

Strength, needs and resilience:
Methods

Brainstorming

1. Problem Statement
   Understand, Identify, Reform

2. Plan, Design
   Who, How, Where, What

3. Implement
   Field Work

4. Analyze & Report Publication

CONSULT
   Existing work, type

IMPACT
   Reform policies, change practices

REINFORCE
   Capacity, research design, advocacy

PARTICIPATE
   Joint analysis, interpretation

INCLUSIVE & PARTICIPATORY
   Design of instrument, methodology

COLLABORATE
   Data collection
SUMMARY
KEY FACTORS:

• CONSTANT ADAPTATION
  • SAMPLING
  • CONTENT
  • LOCALIZATION

EXAMPLE

Elections

Voted in the December 2018 elections (% yes)

Believe the electoral process was credible (% yes)

Why not?
  • Delayed / not organized
  • Not interested
  • Useless
  • Voting station too far

Why not?
  • Voting delayed / canceled in some areas
  • False votes
  • Counting not transparent
  • Cheating during count
Carte 1: Accès aux services et besoins de base par quartier

EXAMPLE
Access to services - Goma

Accès à l’électricité
(% oui – Goma : 46%)

Accès à l’eau potable
(% bon – très bon – Goma : 36%)

Accès aux services de santé
(% bon – très bon – Goma : 29%)

Accès à l’éducation
(% bon – très bon – Goma : 23%)

GLOBAL KEY INDICATORS

EXAMPLE
Core – security, peace, justice…
**EXAMPLE**

**Effectiveness of peacekeeping?**

Local Trust in UN Peacekeeping Operations: Survey Evidence from DRCongo*

Patrick Vina,1 Thomas O’Malley and Pham Nguyen2

This date: November 8, 2019

Abstract

Peacekeeping operations rely on the support of the local population to gather information and to wield power. But local populations may have varying reactions with differing levels of trust. This survey, conducted in the Democratic Republic of Congo, seeks to understand the nature of local trust and its relationship with peacekeeping effectiveness. It examines the perceptions of local residents about the effectiveness of peacekeeping operations, particularly in areas where security is a concern. The findings suggest that local trust is an important factor in the success or failure of peacekeeping missions, highlighting the need for more research in this area.

Keywords: Peacekeeping, Trust, Democratic Republic of Congo, United Nations

Preparation Date: Please do not write in this section

---

**EXAMPLE**

**Ebola**

Institutional trust and misinformation in the response to the 2018-19 Ebola outbreak in North Kivu, DR Congo: a population-based survey

Najma Ahmed1,2,3,4,5,6,7,8,9,10,11,12 and Daniel Campbell2,3,4,5,6,7,8,9,10,11,12

This date: November 8, 2019

Abstract

The Ebola outbreak in North Kivu, DR Congo in 2018-19 was the largest outbreak in the Democratic Republic of Congo, with over 8,000 cases and 4,000 deaths. The response to the outbreak was hindered by misinformation, mistrust, and lack of information. This study examines the institutional trust of communities and the role of media in the response to the outbreak.

Keywords: Ebola, Trust, Media, Misinformation

Preparation Date: Please do not write in this section

---

**Table 1: Trust in states and institutions**

<table>
<thead>
<tr>
<th></th>
<th>Unweighted (N)</th>
<th>Weighted (N:95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>951</td>
<td>-</td>
</tr>
<tr>
<td>Government trust</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trust local authorities</td>
<td>349</td>
<td>32.3% (27.6-37.1)</td>
</tr>
<tr>
<td>Trust provincial authorities</td>
<td>310</td>
<td>32.5% (27.8-37.3)</td>
</tr>
<tr>
<td>Trust national authorities</td>
<td>259</td>
<td>25.4% (20.7-30.7)</td>
</tr>
<tr>
<td>Ebola-related trust</td>
<td>483</td>
<td>40.8% (36.8-44.8)</td>
</tr>
<tr>
<td>Trust government for Ebola response</td>
<td>483</td>
<td>40.8% (36.8-44.8)</td>
</tr>
</tbody>
</table>

---

**Public opinion?**

• **MIXED EVIDENCE—MACRO LEVEL - +, MICRO LEVEL – LOCALLY PROBLEMATIC**

**Unique org. nature**

• **INTERNATIONAL, NOT LIKE LOCAL POLITICAL ACTORS**
RESULTS
Effectiveness of peacekeeping?

- Contact COBs
  - Higher effect for COBs sample
- Effects of knowledge / information
- Frequency of interaction
- Base closure
- Violence
  - (lack variance in COB sample)
    - Some v. none – increase trust
    - High v. some – lower trust
  - Witness - mixed

EXAMPLE
Security and ethnic relations
So…
Is Eastern DRC progressing toward peace?

“Oh, if only it were so simple.”
“I’ll pause for a moment so you can let this information sink in.”

“How can I trust your information when you’re using such outdated technology?”
"I knew we didn’t accomplish anything, but that’s what meetings are for."
D) SLIDES FROM VIRTUAL WORKSHOP ‘ARTIFICIAL INTELLIGENCE (AI), NATURAL LANGUAGE PROCESSING (NLP), AND TEXT-AS-DATA’, PRESENTED BY DR. MARGARET J. FOSTER ON NOVEMBER 16, 2022

AI and NLP for UN DPO Data

Dr. Margaret J. Foster
University of North Carolina, Chapel Hill
margaret.foster@unc.edu
Overview

Part I: Technical Overview
Part II: Promises and Pitfalls in the A4P(+) Context
Part III: Techniques for Analysis and Presentation
Part IV: Overview of sources for additional data

Technical Overview: Terminology

**Machine Learning:**
Formally: Computer systems that can learn and adapt without explicit instructions
Informally: Pattern-finding

**Artificial Intelligence (AI):**
Computers implement utility functions to make decisions

**Big Data:**
Data arriving in great velocity and with great volume; often in complex forms.
Technical Overview: What is NLP/Text–as–Data?

Natural Language Processing (NLP):
Using linguistics, computer science, and AI to allow computer programs to understand human language as spoken and written.

Turns texts (and images!) into data

How?

How does NLP/Text/Images as Data Work?

<table>
<thead>
<tr>
<th></th>
<th>carried</th>
<th>offens</th>
<th>attac</th>
<th>descrp</th>
<th>district</th>
<th>fierce</th>
<th>fight</th>
<th>fire</th>
<th>targ</th>
</tr>
</thead>
<tbody>
<tr>
<td>doc1</td>
<td>0.000</td>
<td>0.053</td>
<td>0.0215</td>
<td>0.1</td>
<td>0.053</td>
<td>0.001</td>
<td>0.032</td>
<td>0.063</td>
<td>0.042</td>
</tr>
<tr>
<td>doc2</td>
<td>0.000</td>
<td>0.000</td>
<td>0.01</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>doc3</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>doc4</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>doc5</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>doc6</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>
Opportunities: Seeing the Bigger Picture Faster

- ML models can summarize themes at scale
- Good for systematic perspective on broader context
- Helps quantify trends and covariate relationships
- Efficiently access historical information
- Identifies new content embedded in reports

Opportunities: Example

Possible Use Cases: automatically code event participants over time

Report of the Secretary-General on the situation in Mali (9/2014)

4. On 15 and 16 June, the Ministers for Foreign Affairs and Ambassadors of Algeria, Burkina Faso, Chad, Mali, Mauritania and the Niger met in Algiers, affording my Special Representative the opportunity to reiterate the call of the Council for proper coordination among international actors in support of the Malian peace process.

Color key:
- Whom
- From where
Opportunities: Example II

Possible Use Cases: extract conflict events

Report of the Secretary-General on the situation in Mali (9/2014)

19. The breaches of the ceasefire increased security risks for civilians and resulted in human rights violations, in breach of article 10 of the preliminary agreement. After MNLA took the town of Aguelhok on 21 May, 66 Imghad Tuareg civilians men, fearing for their security, sought refuge in the MINUSMA camp. They stayed there until 3 July, when MINUSMA negotiated their safe return with MNLA and HCUA. The clashes that took place from 11 to 26 July in Anefis and Tabankort resulted in the killing of at least 4 civilians and the forced displacement of 56 women and 72 children. On 18 August six men were severely beaten by MNLA in Méraka (Gao region) for wearing t-shirts bearing the colours of the Malian flag. In Lerneb, MAA (Coordination) has curtailed the movements of members of the Arab community suspected of being MAA (Platform) sympathizers.

Challenges

Challenge: NLP/Text-as-Data models content

...so must consider omissions, biases, and frames

Report of the Secretary-General on the situation in Mali (12/2014)

16. On 6 November, the Malian defence and security forces took control of the southern bank of the Niger River in Didi, east of Timbuktu, while MAA (Coordination) and MNLA remained in control of the northern bank. In mid-November, the Coordination took control of Zarho (100 km east of Didi), and on 1 December the Platform seized Bamba (30 km east of Zarho). There was no significant violence recorded during those movements.

17. Extremist groups were suspected of killing 16 peacekeepers and injuring 14 others during the reporting period. Improvised explosive devices and anti-vehicle mines placed along routes used by MINUSMA severely hindered its operations. On
A non-exhaustive list of possible sources of bias:

Level: Data Generating Process
Frame: Positivity/negativity biases
Missing not-at-random
Omission
Noise

Level: Analysis
Unit of analysis mismatch
Model vs data generating process
Opacity of decision criteria
Algorithmic biases

Illustration: missingness/noise not at random

Toy example: a wedding invitation for one

Takeaway
Understanding what you see/don’t see requires domain knowledge
Challenges III

**Streetlight Effect**: observational bias of searching where it is easiest to look

In the context of computational methods:
Data-hungry computational tools revolutionize analysis, in "bright" areas

**Takeaway**: The existence of computational tools doesn't imply that the necessary data

- Exists
- Is accessible
- Can be generalized

Challenges: Conclusion

ML/AI are math
Math is a model
Overcoming the Challenges

**Strategies:**
- Evaluate with expertise
- Triangulate with additional data inputs
- Treat ML outputs as tools and not answers

ML as a Tool: Thematic Summaries

ML tools can identify and track themes

**Model:**
Structural Topic Model

**Data:**
Committee on Trade and Development
World Trade Organization

**Visualization:**
STM Insights

Prevalence of "LDC preferences" theme
ML as a Tool: Representative Passages

ML tools can help sift large quantities of data

Task: take 8,500+ paragraphs, find 100 most strongly related to identified theme

Model: Structural Topic Model

Data: Committee on Trade and Development World Trade Organization

Visualization: STM Insights

ML as a Tool: Sentiment in Different Settings

ML can be used to summarize sentiment(*)

Results grouped as desired (eg: country, year, mission, participants)

Model: Sentiment analysis

Data: Shakespeare tragedies

Visualization: Tidytext

ML Tools for Texts

Some other useful tools

- Sentiment/Frequency analysis (https://github.com/juliasilge/tidytext)
- Keyword-assisted topic model (https://keyatm.github.io/keyATM/)
- Word and language embedding (https://huggingface.co/docs/transformers/model_doc/bert)

Additional Data Sources for UN/A4P (+): Text

The UNSC Debates Corpus (Schoenfeld et al. 2021)
What: 82,000+ speeches from 5,748 UNSC Meetings (1995-2020)
Codebook + data available via Harvard Dataverse:
https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/KGYSYH

UN General Debate Corpus (Baturo, Dasandi, and Mikhailov 2017)
What: 7300+ country statements at UNGA (1970-2014)
Visualization tool: http://ungc.smikhaylov.net/
Data available via Harvard Dataverse:
https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/0TJX8Y

PeaceKeeping Operations Corpus (Amicarelli and Di Salvatore 2021)
1,455 reports covering 68 UN Peacekeeping Missions (1994-2020)
Data is available via: https://www.prio.org/jpr/replicationdata

Sources compiled by Evgenja Kroeker
Additional Data Sources for A4P(+): Event Data I

**Peacekeeping Specific:**

International Peace Institute (IPI) Data; monthly mission personnel summaries

Peacekeeping Mandates (PEMA) Dataset; UNSC resolutions

Geocoded Peacekeeping Operations (Geo-PKO) Dataset; Mission deployment maps; UNSG mission progress reports

UCDP Peacemakers at Risk (PAR) Dataset; Factiva (News archive database); UN, NGO and open-source reports

UN Peace Initiatives (UNPi) Dataset; Repertoire of the Practices of the Security Council, UNGA Yearly Reports

Sources compiled by: Evgenija Kroeker, Malina Molska & Andrea Ruggeri

Additional Data Sources for A4P(+): Event Data II

**General Conflict Events**

Armed Conflict Location & Event Data Project (ACLED)

Uppsala Conflict Data Program (UCDP)

Social conflict analysis dataset (SCAD)

Global terrorism database (GTD)

Sexual Violence in Armed Conflict (SVAC) Dataset

Varieties of Democracy (V-DEM)

Sources compiled by: Evgenija Kroeker, Malina Molska & Andrea Ruggeri
Questions?

Questions? Comments?
Contact: margaret.foster@unc.edu
E) REFERENCES


