



Semi-Automated Coding for Conflict Mediation Research: Database Development

Elisa D'Amico



PeaceRep

Peace and Conflict
Resolution Evidence
Platform

PEACE ANALYTICS SERIES



THE UNIVERSITY
of EDINBURGH



University of
St Andrews

Author: Elisa D'Amico

PeaceRep: The Peace and Conflict Resolution Evidence Platform
School of Law, Old College, The University of Edinburgh
South Bridge, Edinburgh EH8 9YL

Tel. +44 (0)131 651 4566

Fax. +44 (0)131 650 2005

E-mail: peacerep@ed.ac.uk

[PeaceRep.org](https://www.peacerep.org)

✉ [@Peace_Rep_](https://twitter.com/Peace_Rep_)

[in https://www.linkedin.com/company/peacerep/](https://www.linkedin.com/company/peacerep/)

Developed within the framework of the Global Fragmentation and Peacemaking Initiative, this report aims to bolster the project's research efforts and serve as a concise manual for semi-automated coding in the broader field of conflict mediation studies.

This research is supported by the Peace and Conflict Resolution Evidence Platform (PeaceRep), funded by UK International Development from the UK government. However, the views expressed are those of the authors and do not necessarily reflect the UK government's official policies. Any use of this work should acknowledge the authors and the Peace and Conflict Resolution Evidence Platform.

Author bio:

Elisa D'Amico is a Post-Doctoral Research Fellow at the University of St Andrews with PeaceRep. Her research focuses on semi-automated methods and web scraping for Conflict Mediation Events, contributing to the Global Peace Actors Database (GLO-PAD). Additionally, she researches shifts in global fragmentation within conflict and peacemaking spheres and examines how external shocks influence mediation processes and outcomes. Elisa holds a PhD in Political Science from Loyola University Chicago.

Acknowledgements:

Thank you to Niamh Henry, Alice Raymond, Allyson Doby, and Mateja Peter for their contributions to the production of this work.

Cover images: Getty Images. All images may be subject to copyright. ©2024

Design: Smith Design Agency

Contents

Executive Summary	1
Introduction	2
1 Background	3
Overview of Conflict Mediation Data	3
Significance and Purpose of Semi-Automation	5
Scope and Objectives	7
2 UCDP Conflicts since 2010	10
The Changing Liberal International Order	11
Selection Criteria and Dataset Description	12
Temporal Constraints and Conflict Duration	13
3 Machine Learning Approach	15
Web Crawling for Conflict Mediation References	16
Data Collection and Cleaning	17
4 Challenges and Solutions	20
Under-specification and Data Noise	20
Handling Large Observations: Issues and Solutions	20
5 Overview of Gathered Data	22
Trends and Patterns in Conflict Mediation References	22
Insights from Semi-Automated Coding	27
Ethical Considerations	28
Bias and Fairness in Data Processing	29
6 Conclusions & Recommendations	31
Future Development and Iterative Review	31
Implications for Conflict Mediation Research	32
Closing Thoughts	32
References	34



Executive Summary

This project leverages recent advances in computational tools and machine learning to automate the analysis of vast quantities of text data related to conflict mediation. By automating the process of identifying, categorizing, and extracting relevant information from diverse online sources, this approach overcomes the limitations of manual coding methods. This allows researchers to access and analyse a much wider range of data, including sources that were previously difficult or impossible to comprehensively access manually.



Introduction

This report, and related database building endeavour, delves into the evolving dynamics of mediation efforts, and third-party involvement. This exploration is motivated by the shifting global norms, with the ongoing trend towards global fragmentation adding a layer of complexity to the understanding of conflict resolution and peacebuilding efforts (Peter, 2023). The imperative to navigate this evolving landscape is the backdrop for the development of our novel mediation database and is discussed in Section 1 of this report. The subsequent sections detail the complexities of semi-automating these phenomena, framed within the broader context of global fragmentation.

Building the foundation for semi-automating the analysis, this report provides an explanation of our web-scraping process, emphasizing its evolution from data science to a multidisciplinary tool. Section 2 explores our machine learning approach, offering an innovative tool for semi-automated analysis. This approach becomes not just a technological advancement but a necessity to adapt to the evolving challenges of conflict mediation in the era of global fragmentation.

The section further focuses on the meticulous methodology of web crawling for conflict mediation references, focusing on the challenges faced in data collection and implementing nuanced mediation efforts with ever-changing scopes. The meticulous process of data collection and cleaning is discussed, as are the nuances of conducting keyword searches within translated reports.

Section 3 delves into the selection criteria and dataset description, providing insights into the temporal constraints and conflict duration considerations within Uppsala Conflict Data Program (UCDP) conflicts since 1988 (Davies and Engström, 2024). Section 4 addresses challenges and solutions in our web scraping strategies. It examines under-specification, data noise, and issues with handling large observations, highlighting the need for adaptable and precise conflict analysis methods.

Ethical considerations, including bias and fairness in data processing, are explored in Section 5, examining the responsible use of web-scraped information. The section further provides an in-depth analysis of trends and patterns in conflict mediation, reflecting the impact of global events and normative shifts.

The concluding section offers practical suggestions to enhance semi-automated methodologies in mediation data. It captures the transformative potential of semi-automating conflict mediation research, shaping the landscape of peace and conflict studies in the era of global fragmentation.



1 Background

The current landscape of mediation data sources, including notable ones such as the Uppsala Conflict Data Program (UCDP) (Davies and Engström, 2024), International Crisis Group (ICG) (2024), and African Peace Processes (APP) (Duursma and Gamez, 2023), has proven insufficient in capturing the nuanced dynamics of mediation within the context of emergent global fragmentation. These databases, while valuable, often struggle to keep pace with the rapidly evolving nature of conflicts and the intricacies of modern geopolitical shifts. The limitations in precision, scope, and potential biases inherent in these data underscore the need for a more sophisticated and timely approach.

The advent of advanced data science, coding, and automation technologies presents a transformative opportunity. By harnessing these tools, we can now capture the multifaceted dimensions of conflict mediation with unprecedented accuracy, breadth, and adaptability. This realization has driven the initiative to develop PeaceRep's new global fragmentation mediation database, which seeks to leverage cutting-edge technologies to offer a comprehensive and dynamic understanding of mediation processes in an era marked by rapid global changes and challenges.

Overview of Conflict Mediation Data

The existing data on conflict mediation faces a myriad of challenges that significantly limit its utility and accuracy. One of the primary issues is the lack of data accuracy and completeness. Data collection in conflict mediation is often challenging due to the complexity and sensitivity of the situations involved. Inaccuracies may arise from misinformation, lack of transparency, or intentional manipulation of information by conflicting and third parties. Moreover, the incomplete reporting of mediation efforts can lead to an inadequate understanding of the dynamics at play. Additionally, the passage of time erodes the reliability of data sources, as websites change or disappear, documents are lost, and institutional knowledge fades. This temporal decay not only reduces the availability of crucial information but also complicates the verification of historical data, further undermining the integrity of the analysis.

Regional and temporal coverage is another critical concern. Many conflict mediation datasets exhibit biases toward certain regions or time periods, reflecting historical and geopolitical priorities. This bias limits the generalizability of findings and can skew analyses, particularly when attempting to draw conclusions about the effectiveness of mediation strategies across diverse contexts.



Definition and coding differences pose significant challenges to data interpretation. Various databases employ different definitions and coding criteria for key concepts, such as identifying third-party actors and characterizing the broader peace processes involved. Furthermore, while many data collection efforts acknowledge the presence of external mediation, they often fail to specify who the mediators are, how many there are, and their characteristics. This lack of standardized terminology and detail not only hampers comparability but also complicates efforts to aggregate data for cross-case comparisons.

Quality of data is another critical aspect. The reliability and depth of information on the mediation process, including the quality of negotiations, the implementation of agreements, and the sustainability of resolutions, are often lacking. Without a comprehensive understanding of these factors, it is challenging to assess the true impact and success of mediation initiatives.

Beyond these issues, the existing data fails to capture the changing nature of mediation in the current landscape of global fragmentation. The scope of various conflict mediations is expanding, encompassing economic and humanitarian dimensions. Economic interests and humanitarian concerns are increasingly influencing the mediation agenda, adding new layers of complexity that traditional datasets may not adequately capture.

Moreover, the potential biases of the mediators are not sufficiently addressed in existing datasets. Mediators may bring their own geopolitical interests, cultural biases, or preconceived notions into the mediation process, influencing the outcomes. Understanding and accounting for these biases is crucial for a nuanced analysis of mediation effectiveness. Finally, there is a broader trend of ongoing large-scale mediation initiatives, such as the Berlin Process or AU-led mediation, that significantly influence conflict dynamics. However, when we analyse single mediation events in isolation, we fail to recognize their importance within the context of these larger processes. This lack of consideration complicates our understanding of how mediators operate both within large peace processes and outside of them. These differences are likely crucial for assessing long-term mediation effectiveness.



In conclusion, while existing data on conflict mediation is valuable, it is essential to recognize and address these limitations to enhance its reliability and relevance. Standardizing definitions, improving data accuracy and completeness, expanding regional and temporal coverage, promoting interoperability, and considering the evolving nature of mediation dynamics are crucial steps toward building a more comprehensive and insightful understanding of conflict resolution efforts on a global scale.

Significance and Purpose of Semi-Automation

In addressing the challenges associated with existing data on conflict mediation, our approach emphasizes the significance and purpose of semi-automation in building a more robust and nuanced database. Semi-automation represents a strategic fusion of human expertise and machine learning techniques, offering unprecedented advantages in data capture, reliability, and scalability.

- **Nuanced Data Capture:** By employing semi-automation techniques, we enhance the precision and depth of data capture in conflict mediation processes. Unlike traditional manual coding, which may be prone to biases and inconsistencies, our approach leverages machine learning algorithms to identify and categorize nuanced information. This ensures a more accurate representation of mediation events, capturing subtleties that may be overlooked in manual coding.
- **Translingual Data Extraction:** One of the groundbreaking features of our methodology is the ability to pull data in a translingual way. Machine learning algorithms enable the extraction and interpretation of information from diverse languages, overcoming a significant limitation in traditional manual coding where language barriers often hinder comprehensive data collection. This translingual capability broadens the scope of our database, providing a more inclusive and global perspective on conflict mediation efforts.
- **Reliability at Scale:** The use of semi-automation allows for the systematic processing of vast amounts of data, offering a level of reliability and consistency that manual coders struggle to achieve at such scale. Through machine learning algorithms, we can analyse and code large datasets efficiently, ensuring a more comprehensive understanding of the mediation landscape across various regions and time periods.



- **Web Crawls and Information Accessibility:** Our methodology incorporates web crawls, an advanced technique in the information age, enabling the systematic exploration and collection of data from online sources. We draw from a diverse array of platforms, including reputable news outlets—both local and national—NGOs, IGOs, and respected institutional websites (e.g. [ReliefWeb](#)). This approach facilitates the identification and coding of smaller scale mediation events within broader processes, providing a level of granularity that is increasingly vital in understanding the intricate dynamics of conflicts. Web crawls contribute to the accessibility of information in an era where online sources play a pivotal role in shaping narratives and disseminating data.
- **Adapting to Changing Dynamics:** In doing the above, the database will chronicle the evolving landscape of mediation efforts, examining who gets involved, when, and why. By considering the changing nature of mediation attempts, we can identify patterns and factors that influence the success or failure of these initiatives. This adaptability is crucial in understanding the dynamic context of global fragmentation.
- **Providing a Complete Depiction:** By integrating all these aspects, our database seeks to offer a more complete depiction of the evolving nature of mediation efforts. It considers the interplay between global fragmentation, changing mediators, biases, and the diverse scope of conflicts. This holistic approach contributes to a more nuanced understanding of how mediation strategies adapt to the complex challenges of the contemporary geopolitical landscape.
- **Sustainability:** The sustainability of our methodology is a key advantage, as once the web crawling system is established, it can be reused periodically to collect new data with minimal additional costs. While the initial setup requires a significant investment of time and expertise, this upfront effort is offset by the long-term benefits of the system's reusability. By automating the source collection process, we can efficiently gather updated information on a regular basis without the need for extensive manual intervention, ensuring that our dataset remains current and relevant over time. This sustainable approach not only maximizes resource efficiency but also enhances our capacity for ongoing analysis and monitoring of trends in conflict mediation.



In conclusion, the significance and purpose of semi-automation in building our conflict mediation database lie in its ability to enhance data capture, ensure reliability at scale, transcend language barriers, utilize web crawls for nuanced information, and adapt to the dynamic nature of contemporary conflicts. Through the integration of human expertise and machine learning capabilities, our approach represents a powerful tool for researchers and policymakers seeking a more comprehensive and nuanced understanding of conflict mediation in the modern world.

Scope and Objectives

Scope: Our primary goal is to construct a comprehensive mediation database that spans the recent era of Global Fragmentation, dating back to 2010. This initiative aims to capture all existing conflicts which exceed 1,000 battle deaths per year, as categorized by the Uppsala Conflict Data Program (UCDP) during this period (Davies and Engström, 2024). The scope of our work is ambitious, seeking to provide a holistic understanding of conflict mediation on a global scale.

Objectives:

■ Capturing Ongoing Conflict Processes:

- Our database will go beyond traditional categorizations, capturing not only the broader nature of ongoing conflict processes but also the intricate sub-mediation events occurring within them.
- This depth of data will allow for a nuanced analysis of mediation efforts throughout the lifespan of conflicts, providing insights into the dynamics that influence their trajectories.

■ Understanding Actor Involvement in Mediation Efforts:

- An objective of this database is to comprehensively document the various actors engaged in mediation, including countries, organizations, and individuals across different contexts and types of mediation.



- By capturing specific mediation events - whether they lead to formal agreements or result in no agreement at all - we aim to provide a more thorough account of mediation activities. This information holds significant value for a broad audience, including practitioners and policymakers who may not engage with political science literature, thereby enhancing the accessibility and practical relevance of our findings.
- This descriptive data, which is currently lacking in detail and scale in existing sources, is essential for painting an accurate picture of the mediation landscape.

■ Understanding Scope:

- We aim to categorize mediation events based on their scope, distinguishing between political, humanitarian, and economic dimensions. These categorizations are preliminary and will evolve based on Natural Language Processing (NLP) methods, Topic Modelling, and clustering techniques.
- This approach recognizes the multifaceted nature of conflicts and enables us to predict outcomes more accurately by understanding the varied influences at play during the resolution process.
- By incorporating the nature of mediation events and their scope, our database will contribute to understanding and predicting the outcome of conflict resolution more effectively. Understanding whether the mediation is driven by political considerations, humanitarian concerns, or economic interests provides crucial insights into the potential success or challenges in achieving sustainable peace.

■ Analysing Mediator Bias:

- Our work will delve into the nature of mediator bias, focusing on the perspectives and leanings of mediators towards the warring parties using advanced network analysis.
- This analysis is pivotal in predicting the outcomes of peace processes, as mediator bias can significantly influence negotiations and ultimately, the implementation of agreements.



■ Adapting to Changing Dynamics:

- In doing the above, the database will chronicle the evolving landscape of mediation efforts, examining who gets involved, where, when, and why.
- By considering the changing nature of mediation attempts, we can identify patterns and factors that influence the success or failure of these initiatives. This adaptability is crucial in understanding the dynamic context of global fragmentation.

In summary, our scope and objectives centre around creating a dynamic, comprehensive mediation database that not only captures the diversity of conflicts and mediation events but also analyses their scope, examines mediator biases, and can be leveraged to understand the changing dynamics of global fragmentation. This initiative represents a significant step toward unravelling the complexities of conflict resolution in the recent era of peacemaking.



2 UCDP Conflicts since 2010

The choice of data selection from the Uppsala Conflict Data Program (UCDP) stems from its systematic and consistent approach to measuring various forms of conflict (Davies and Engström, 2024). The UCDP provides a comprehensive database that covers state-based armed conflict, non-state conflict, and one-sided violence. These categories are mutually exclusive, allowing for a clear distinction between different forms of organized violence. The use of a common intensity cut-off, such as 1,000 fatalities in a calendar year, ensures a standardized threshold for inclusion, facilitating comparability across different conflicts.

One of the strengths of UCDP data is its ability to aggregate diverse forms of conflict into a broader category of "organized violence." This approach recognizes that conflicts can take different forms and involve various actors, allowing researchers to analyse patterns and trends in violence more holistically.

In the post-Cold War era, the nature of conflicts and conflict mediation underwent significant changes. The United Nations (UN) played a crucial role in promoting decolonization and became more actively involved in conflict mediation and resolution. The end of the Cold War led to a shift in the dynamics of international relations, creating new opportunities and challenges for conflict resolution efforts. As superpower rivalry diminished, regional conflicts gained prominence, and the UN stepped in to mediate and prevent the escalation of violence (Belhadj Klaz & Mariani, 2022; Eppe et al., 2024; Magara, 2023; Nash, 2021; Nash et al., 2022; Freer, 2022; Houghton, 2024).

These changes are even more drastic in the current era of Global Fragmentation. The UN's involvement in conflict mediation became more pronounced through peacekeeping missions, diplomatic initiatives, and efforts to address the root causes of conflicts. The organization's peacekeeping operations aimed to maintain or restore peace in areas affected by conflict, providing a framework for negotiations and reconciliation.

By focusing on UCDP conflicts from 2010 onwards, we can capture this critical period of change in the international landscape. This time-frame allows for the examination of conflicts in this era of changing peacebuilding, where we are seeing the emergence of more regional and "non-Western" actors in mediation processes. Analysing conflicts during this period provides insights into the effectiveness of UN interventions as opposed to other forms of mediation, the evolution of conflict dynamics, and the challenges faced in mediating diverse and complex disputes.



In summary, the choice of UCDP data from 2010 onwards is justified by its consistent and comprehensive measurement of organized violence, covering state-based armed conflicts, non-state conflicts, and one-sided violence. Additionally, this period aligns with significant shifts in global politics, where the UN emerged as a key player in conflict mediation and peacekeeping, making it a valuable dataset for understanding modern conflict and conflict resolution dynamics.

The Changing Liberal International Order

The decision to examine conflicts from 2010 onwards is informed not only by the systematic approach of the Uppsala Conflict Data Program (UCDP) but also by the shifting dynamics in global politics, specifically characterized by the pragmatic turn within collective responses (Davies and Engström, 2024). This pragmatic turn reflects a divergence between doctrinal principles and practical implementation, particularly evident within the Liberal International Order (LIO) (Peter, 2023).

The pragmatic turn, characterized by a de-prioritization of human protection within collective instruments, stems from tensions within assertive liberal internationalism. This shift is evident in the compromise of the principled nature of UN peacekeeping, especially with the incorporation of counter-terrorism efforts and stabilization mandates. This broader context influences our data selection. By focusing on conflicts from 2010 onwards, we capture a time when conflicts became more diverse and involved a spectrum of actors, both illiberal and liberal, in increasingly disparate mediation efforts. The pragmatic turn not only affects the nature of conflicts but also introduces biases and varying scopes among mediators and mediation efforts, calling for our novel approach.

Incorporating conflicts from this period into our analysis allows us to observe how these pragmatic shifts within the LIO impact the data landscape. The inclusion of diverse actors and the nuances of their mediation approaches highlight the challenges and complexities inherent in conflict resolution during this transformative era. Therefore, our data selection is not merely chronological but is strategically aligned with the crucial juncture when pragmatic considerations began shaping the dynamics of conflict mediation, influencing the roles of different actors and their biases, as reflected in our dataset.



Selection Criteria and Dataset Description

The dataset encompasses a comprehensive set of information crucial for understanding conflicts, mediation efforts, and their outcomes. It includes details on the geographical context of conflicts, beyond UCDP conflicts, further specifying countries and regions mentioned in UN Security Council debates (Davies and Engström, 2024). Unique identifiers for conflicts and dyads allow for precise tracking, while actor information delineates the parties involved, including governments, rebel groups, and supporting actors. The dataset classifies conflicts into different types, such as interstate, intrastate, and internationalized (according to UCDP classification), providing insights into the nature of the disputes. Temporal aspects, like conflict start and end dates, enable the analysis of conflict duration related to mediation efforts and processes.

Additionally, the dataset incorporates a robust section on mediation efforts, capturing details on broad mediation processes, singular events, and links to previous efforts. It identifies the location, start date, and leading actors in mediation, offering insights into the dynamics of negotiations. The inclusion of other third-parties, the number of mediators, and gender considerations provide a nuanced understanding of the mediation landscape. Furthermore, variables on mediator power and types of mediation initiatives shed light on the influence and approaches of those facilitating the peace process.

The dataset also encompasses the scope of mediation efforts, including humanitarian, economic, and political dimensions. It highlights whether negotiations address the entirety of the conflict agenda, contributing to an understanding of the comprehensiveness of mediation outcomes. It also allows the user to parse the success of different types of mediation events including different types of mediators. Activities related to good offices, negotiations, hosting and chairing, and coordinating third-parties are detailed, offering a nuanced view of the methods employed in the mediation process. The dataset further captures the names of agreements resulting from mediation efforts and identifies third-parties that signed or witnessed these agreements providing interoperability to other datasets, such as PA-X Peace Agreements Database (Bell et al., 2024).



In its initial iteration, the dataset provides a robust foundation for understanding conflicts and mediation efforts, offering a detailed account of geographical, temporal, and actor-related variables. However, recognizing the inherent complexities and potential biases in mediation processes, a second version of the dataset is envisioned to incorporate advanced machine learning techniques. Specifically, clustered random forest sentiment analysis, among other relevant methods, will be employed to identify and quantify mediator biases. This innovative approach will take into consideration various affinities, including geographical and political factors, to discern subtle nuances in the mediation dynamics. By leveraging machine learning, the second version aims to unveil latent patterns and sentiments within the mediation data, providing a more nuanced understanding of the role of mediators and their potential biases in shaping conflict resolution outcomes. This enhancement reflects a commitment to advancing the dataset's analytical capabilities, making it a more powerful tool for researchers seeking deeper insights into the intricate interplay of conflicts and mediation efforts on a global scale.

For the spatial scope of the dataset, our plan is to initially focus on UCDP conflicts since 2010 in Africa (Davies and Engström, 2024). This targeted approach allows us to create a more region-specific subset of conflicts, ensuring depth as well as breadth, providing deeper insight into the dynamics and mediation efforts.

Once this initial subset is established, it lays the foundation for potential comparative analyses with conflicts from other regions in subsequent iterations. This phased approach provides a targeted and region-specific understanding before broadening the dataset's spatial scope. Overall, by starting with conflicts in Africa, we aim to enhance the dataset's relevance and applicability to the complexities of conflicts and mediation efforts in various contexts.

Temporal Constraints and Conflict Duration

The temporal constraints applied to the dataset centre around the post-2010 period is a deliberate choice, rooted in the recognition that the post-2010 era marks a pivotal juncture characterized by dynamic shifts in global geopolitics and the notable changes in the United Nations' involvement in conflict mediation. This period captures the evolving landscape of conflicts as they transformed from Cold War dynamics to more diverse and region-specific challenges.



Focusing on conflicts post-2010 is strategic, capturing a shift from less detailed, unidirectional conflict resolutions with more coordinated international involvement before this period. The dataset aligns with the complexities of conflicts in this era, allowing a more accurate examination of mediation dynamics. This approach intentionally omits an era marked by less nuanced international involvement in conflict resolution.

The decision to exclusively consider conflicts post-2010 is influenced not just by evolving conflict mediation but also by technological advances since the 2010s. A semi-automated method benefits from the digitized nature of information in this period, enhancing data extraction capabilities. Previous decades lacked comprehensive digitization, limiting systematic data analysis. The post-2010 period, marked by accelerated digitization and globalization, presents a unique opportunity to leverage technology for unprecedented scale and nuance in analysing conflict mediation efforts globally.

Our approach encompasses all stages of mediation, including efforts adjacent to mediation, funding conferences, actual negotiations, good offices, and post-agreement implementation. We ensure that the dataset captures mediation initiatives during all of the critical phases of conflicts, where international intervention and mediation efforts are most pertinent.

This approach provides a clear and concrete metric for structuring the dataset at an aggregate level. This approach enhances the dataset's capacity to illuminate the dynamics of conflict mediation during active, or otherwise ongoing disputes, offering valuable insights into the effectiveness and challenges of international efforts to bring about peace in diverse regions.



3 Machine Learning Approach

Utilizing [Kalev Leetaru's web crawling methodology](#) as a foundation, we employ an innovative machine learning approach, incorporating best practices, to aggregate sources pertaining to third-party mediation in UCDP conflicts since 2010 (Davies and Engström, 2024). Unlike the manual labour-intensive work using platforms like LexisNexis and Factiva, our methodology is all-encompassing. However, given the current under-specified criteria and the real-time translation of machine learning outputs in over 65 languages, the initial data extraction results in millions of observations. Consequently, extensive cleaning and refinement are necessary.

In plain terms, our database employs a combination of Common Crawl and real-time APIs from major news websites to gather comprehensive event data. Common Crawl (2008) is a vast archive that regularly scans the internet, collecting articles and online content. This resource provides us with a historical perspective on global events. Access to this valuable corpus is free. This archive updates approximately monthly. In parallel, we utilize real-time APIs to access the latest news articles directly from systematic sources, ensuring that our data is able to capture recent major events in near-real-time.

To access our "mediation" sources and events, we integrate both crawling and scraping techniques in our data collection process. Crawling involves systematically exploring the web for new content, while scraping extracts specific data from these pages.

To enhance our data quality, we employ entity recognition, which identifies key conflict actors within the articles. This enables us to accurately geolocate "events" as they are reported. Additionally, we utilize natural language processing (NLP) techniques to capture a diverse array of mediation and mediation-related events.

In comparison to traditional human coding, our automated approach significantly reduces the workload associated with data collection. While the cleaning process still requires a considerable effort, it is manageable for a relatively small team, especially when compared with the extensive work required for translation and manual data entry. Tasks that would typically take years to complete can now be accomplished in months, enabling us to produce timely insights from a wealth of data efficiently.



Web Crawling for Conflict Mediation References

First, we pull sources for events where actors engage in mediation. The methodology involves identifying instances where a party explicitly acts as a mediator between two or more parties. The source is attributed to the mediator, and adversaries are designated as the targets. This pull is exclusively applied when a party explicitly meets with others in the role of a mediator.

The project also accounts for "good offices," making a contribution to data on preliminary mediation efforts. Similar to pulling events when actors are engaging in mediation, for negotiations, bargaining, or discussions, we pull data which uniquely captures non-full scale mediation engagement, but rather the preliminary talks. This data is collected when reports explicitly indicate negotiations, bargaining, or discussions during meetings or consultations. This is applicable when the purpose of the meeting involves negotiation, and even terms like "holding talks" and "discussions" are treated as negotiations, considering them as reciprocal events.

By systematically identifying and categorizing these instances, the data extraction process ensures a comprehensive understanding of diplomatic activities, contributing valuable insights to the existing literature.

This methodology is advantageous for several reasons. Firstly, it operates on a vast scale, pooling from over one billion references per year, making it one of the most extensive open access spatio-temporal options available. Furthermore, the methodology's adaptability and scalability make it valuable for studying and navigating the intricate dynamics of global events. In addition, the method's real-time nature ensures its relevance and responsiveness to real-time changes and holding historical observances. This adaptability is crucial for staying abreast of the dynamic nature of global events and trends.



Lastly, again drawing on Kalev Leetaru's methodology, the incorporation of the translingual nature of the web crawl marks a significant leap in real-time news translation and dataset building alike. Every monitored news report is translated in real-time, eliminating language barriers often faced by manual coding teams. The translation channels, incorporating stages like language detection, word segmentation, and morphological analysis, ensures efficient and accurate translation, even despite high levels of information-intake. With dynamic language models and user adaptation, it allows for tailored translations, fostering direct communication between the system and users and eliminating many English-only news source biases.

Data Collection and Cleaning

After the initial source extraction is completed, it yields an extremely large dataset of potential third-party engagements in mediations and negotiations. For example, the initial extraction for Syria in 2023 produced 8,223 potential sources. Due to the broad parameters of the initial pull, there is significant data noise and a number of duplicate events within these sources. To refine the dataset, we apply straightforward keyword recognition to both exclude and include sources based on a progressively refined set of criteria. This filtering process allows us to narrow down relevant events more effectively while systematically reducing unnecessary or redundant data points.

The full data includes each of the UCDP conflicts from 2010-present that are also present on the UNSC agenda (Davies and Engström, 2024). When pulling the full data sources across all conflicts and years, over four million sources were returned. In order to clean some data noise and retain only sources that reference actual mediation or negotiations, we use R to extract the sources and keyword search the data. Much noise is related sources reporting conflict activities with vague reference to the fact that there are ongoing mediations. We use a variety of inclusion and exclusion criteria across the data to refine this problem. To view a sample of the R code used to search this data, please observe the following:



```
# --MEDIATION-- #

mediat <- keyword directory(directory, keyword =
c('mediat', 'peacebuild', 'arbitrat', 'facilitat',
'interven', 'dialogue'), surround lines = 1, full names =
TRUE)

economic <- apply(mediat ,2,as.character) write.
csv(mediat,"mediation.csv")

# --GOOD OFFICES-- #

goodoffice <- keyword directory(directory, keyword =
c('negotiat', 'good office', 'discussion', 'bargain',
'talks'), surround lines = 1, full names = TRUE)

goodoffice <- apply(goodoffice ,2,as.character)
write.csv(goodoffice,"goodoffices.csv")
```

This will provide individual csv/excel sheets to manually read through document reference to certain keywords/concepts for manual coding purposes (rather than reading full documents or individually going into PDFs and keyword searching).

Following the keyword search and extraction of relevant data into CSV files, the next step involved a manual review of the output to identify instances of third-party engagement in mediation or good offices. The CSV files, namely "mediation.csv" and "goodoffices.csv," contained references to documents where keywords associated with conflict resolution were found.



To systematically examine these references, I utilized the reference links provided in the CSV files. The reference links served as direct pathways to the sources, enabling observation of the context and details surrounding each instance of engagement in mediation or good offices. This manual process allowed for a more nuanced understanding of the nature of third-party involvement in conflict resolution activities.

The manual search involved scrutinizing the content of each document to identify key indicators such as the actors involved, the nature of the mediation or negotiation, and the outcomes. Additionally, it provided insights into the broader context of the conflict and the role of third-parties in facilitating resolution and allowed us to consider mediation scope and whether it is linked to a broader process.



4 Challenges and Solutions

In our project, we deliberately cast a wide net during the initial data pull to capture all possible sources related to UCDP conflict countries engaged in talks or mediation (Davies and Engström, 2024). This intentional inclusivity, known as under-specification, is aimed at avoiding the omission of any relevant information. However, this broad approach results in a challenge – data noise. This section discusses how we navigate this challenge and employ a text search feature to filter out irrelevant information and focus on mediation and negotiation events. Additionally, we address the handling of large observations, particularly in dealing with translation issues, emphasizing the ongoing effort required to refine our language translation process and maintain the effectiveness of our data collection.

Under-specification and Data Noise

Under-specification refers to the intentional broadness in the initial data pull, aiming to encompass the entire universe of sources related to UCDP conflict countries engaged in talks or mediation (Davies and Engström, 2024). The primary objective is to avoid missing any potentially relevant information. However, this approach leads to a considerable volume of observations, resulting in a notable amount of data noise.

The strategy involves casting a wide net during the initial data pull to ensure inclusivity. This step is crucial to capture a comprehensive range of references to mediation and negotiation events. This is something that manual coding strategies and focusing on only select sources often leaves out.

To address the challenge of data noise, as highlighted in Section 2, a text search feature is employed. This feature aids in refining the dataset by filtering out irrelevant information and focusing on events related to mediation and negotiation.

Handling Large Observations: Issues and Solutions

Given the nature of our under-specification in the web crawl, the handling of large observations is a critical aspect of our project. The text search feature primarily relies on identifying specific keywords, such as “mediation,” “good offices,” “negotiation,” and so on, within machine-translated and native English documents. This targeted approach helps narrow down the dataset to relevant entries.



One limitation arises from potential translation issues, requiring a meticulous consideration of linguistic nuances. Ensuring accurate translations is vital, as discrepancies may hinder the identification of essential keywords. Continuous efforts are needed to refine the language translation process and identify potential proxy words that convey similar concepts.

Addressing translation challenges is an ongoing process. Regular reviews and updates are necessary to enhance the accuracy of the text search feature. This iterative approach ensures that the system adapts to new linguistic nuances and avoids overlooking critical information due to language variations, particularly as more regions enter the dataset.

Given the evolving nature of language and potential shifts in terminology, it is essential to acknowledge that refining the text search feature and addressing translation challenges is an ongoing effort. Regular updates and improvements will be crucial to maintaining the effectiveness of the data collection and maintenance process.



5 Overview of Gathered Data

Having discussed the methodologies employed for data gathering thus far, we now delve into a detailed exploration of the collected data so far. This section shows the nuances revealed by our meticulous efforts, transitioning from the methods of data gathering to a comprehensive understanding of what the gathered data portrays.

First, we analyse the collected data to reveal trends and patterns in conflict mediation references, showcasing insights shaped by global events and normative shifts. This exploration emphasizes the transformative potential of semi-automated research methods, adapting to challenges posed by global fragmentation.

Expanding on earlier methodologies, we explore the implications of semi-automated coding in conflict mediation. This involves reflecting on technological advancements and the necessity of adapting to evolving challenges in conflict resolution amidst global fragmentation.

Ethical considerations are discussed as we speak to the responsible use of web-scraped information, reflecting on the significance of bias and fairness in data processing. This aligns our methodologies with the principled nature of mediation efforts in the complex global landscape.

A critical examination of potential biases in our data is undertaken, with a focus on strategies for fairness. Precision and adaptability in data processing methodologies are highlighted to uphold research integrity, acknowledging the impact of global fragmentation on conflict mediation dynamics.

This overview aims to transparently convey the complexities of conflict mediation data within the era of global fragmentation, enriching collective knowledge of the challenges and opportunities in this dynamic field, and our ability to leverage big data to achieve this data building endeavour.

Trends and Patterns in Conflict Mediation References

In our data-building initiative, we employed a dual approach, emphasizing depth with a detailed focus on the Sudan/South Sudan conflict since 2010, and breadth by acquiring comprehensive data on all UCDP and UNSC Agenda conflicts for 2024 through web crawling and semi-automated methods.



Through our preliminary analysis on the African context, we have identified emerging trends that we anticipate will become more pronounced within the broader context of African conflicts and potentially extend beyond. Notably, there is a discernible upward trajectory in the number of mediation events over time. This suggests a growing frequency of mediation attempts, indicative of diverse actors and preceding events leading up to the eventual pursuit of peace. Interestingly, the humanitarian scope seems to mirror the overall trend in mediation events whereas the sheer political scope is much more sparse.

Figure 1: Trends in Mediation Events

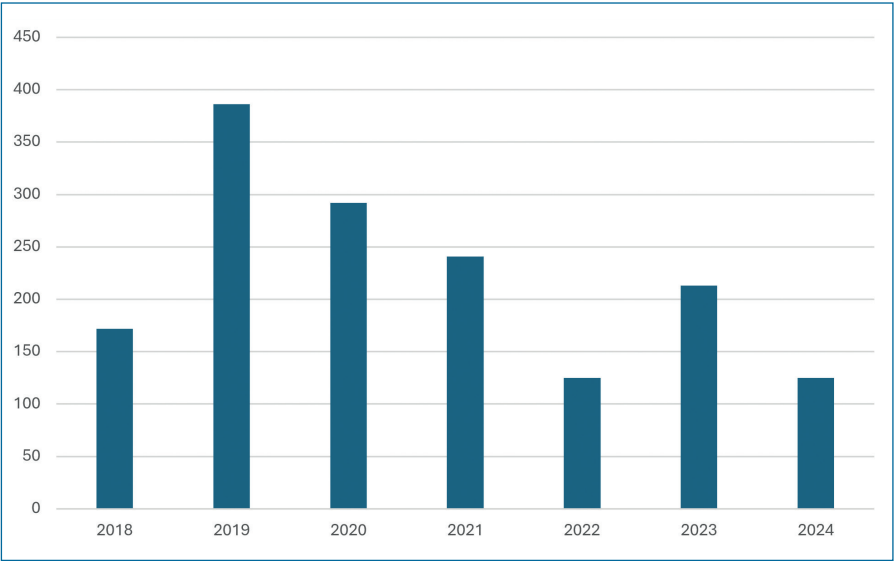
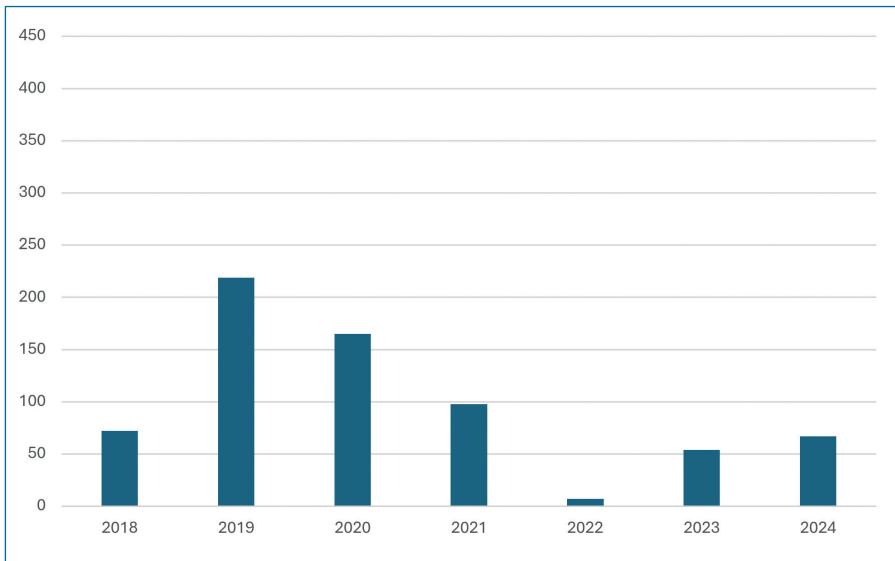




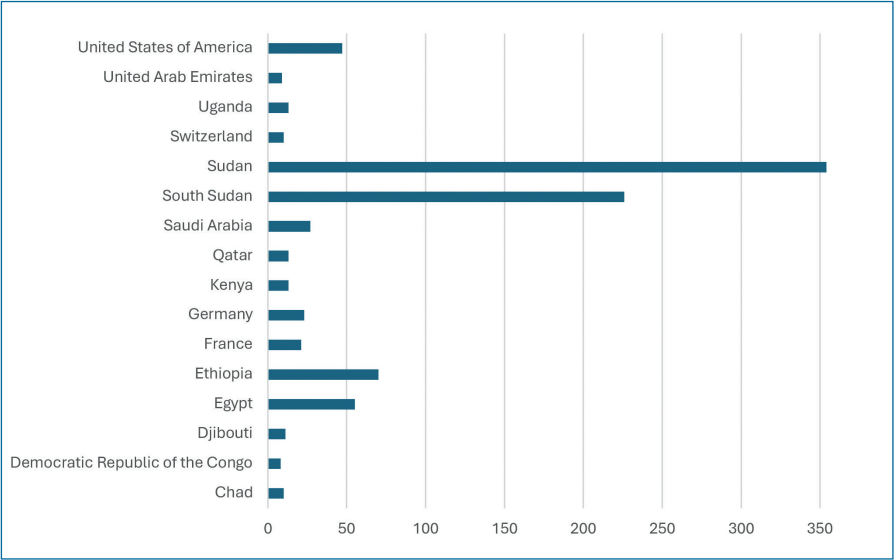
Figure 2: Humanitarian Scope of Mediation Events



Another noteworthy observation is the concentration of mediation events within the geographic locations of ongoing conflicts in 2022, as opposed to third-party countries or locations. This spatial concentration implies a localized focus and a heightened relevance of mediation efforts directly within the conflict zones.



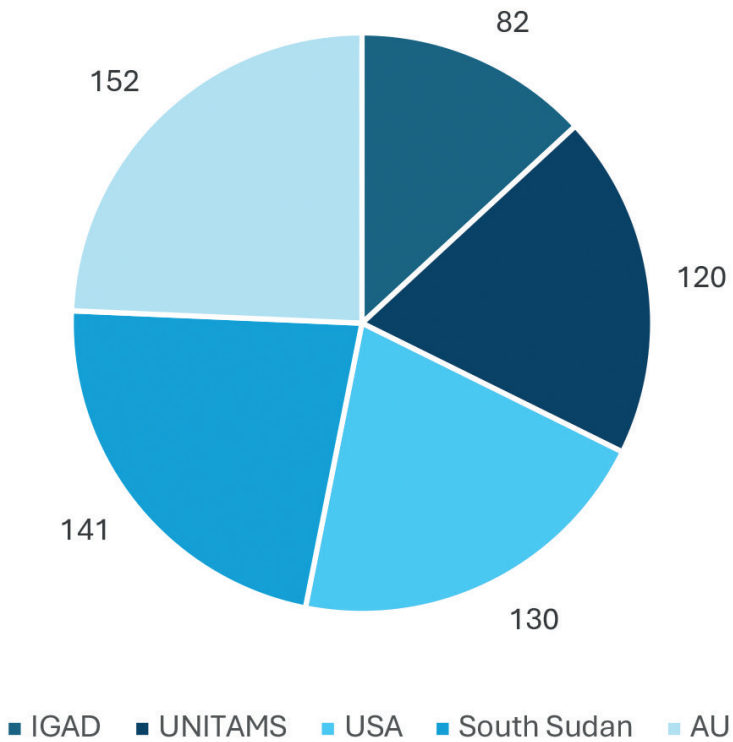
Figure 3: Mediation Event Location



Examining the main mediators chairing these events in 2022, we observe a concentration around key entities such as the African Union, Kenya, and Angola. This concentration indicates a reliance on specific mediators within the African context, underscoring their pivotal role in conflict resolution.



Figure 4: Mediator Distribution



A particularly insightful finding is that only 21.3% of the mediation events in 2022 resulted in a peace agreement. This suggests that ongoing mediation processes involve numerous events that do not necessarily lead to a single solution. Instead, these events require time to come to fruition, highlighting the intricate and protracted nature of conflict resolution.



Insights from Semi-Automated Coding

In transitioning to insights from our semi-automated coding and web-scraping strategy, we uncover the critical role of semi-automated data gathering in capturing the evolving nature of conflict resolution efforts, particularly in the post-2010 period marked by increased internationalization, notably through extensive UN involvement. The semi-automated approach proves to be ideal and comprehensive in its ability to assimilate vast and diverse data sources.

The web crawl component of our project adds a valuable layer to our understanding by pulling in more localized data. This granular perspective offers a clearer and more nuanced picture of mediation efforts, providing insights into the multitude of events, visits, and complexities inherent in broader mediation processes. Moreover, it unveils evidence of sub-national, alternative, and non-mainstream mediation efforts, an aspect that has been overlooked in scholarly work, pointing to a promising avenue for future mediation research.

Particularly striking is the emphasis on the importance of specific mediators and actors, prominently featured in more local news sources. This underscores their perceived significance not only in the broader mediation process but also in shaping the mediation landscape within the region. For instance, Nigerian President Obasanjo and Angolan President Lourenço emerge as key mediators, consistently relied upon in multiple African conflicts, highlighting the regional reliance on certain individuals for mediation roles.

The local news sources not only elevate the importance of individual mediators but also shed light on the intricacies of broader peace processes, often “AU-led peace processes.” This suggests that mediation events may serve a purpose beyond individual conflicts, contributing to a more comprehensive and thorough approach to peace talks and negotiations.

Additionally, our semi-automated coding reveals information about the role of women as informal mediating actors. This contributes to current gaps in our knowledge about the role of women in peace processes as peace actors, rather than merely a party allotted provisions in such processes (Wise, 2023). This underscores the need to explore and understand the role of women in conflict mediation more comprehensively and contributes to ongoing work in PeaceFem.



In conclusion, the insights garnered through semi-automated coding illuminate the evolving landscape of conflict resolution efforts, capturing both the internationalization of mediation and the significance of localized dynamics. The nuanced understanding of specific mediators, sub-national efforts, and broader peace processes, lay a foundation for more informed and targeted research, providing valuable direction for future mediation studies.

Ethical Considerations

In discussing ethical considerations, it is crucial to acknowledge that while our data is sourced from publicly available information, the nature of conflict mediation research necessitates a heightened sensitivity to the information gathered, especially when dealing with details related to rebel groups. The ethical compass of our project pivots on the responsible handling of data, recognizing the potential implications and sensitivities surrounding conflict-related information.

Given the complex and often delicate nature of conflict mediation, privacy concerns and potential risks associated with revealing sensitive information about individuals or groups, particularly rebel factions, are important. Ethical considerations dictate a cautious approach to ensure that our research does not inadvertently contribute to harm or compromise the safety of those involved in conflict or peace building efforts.

Moreover, maintaining clean and accurate data is paramount in upholding ethical standards. In response to this imperative, our team consistently meets and calibrates our coding metrics, fostering a collaborative and meticulous approach to data processing. This ensures not only the reliability of our findings, but also aligns with ethical standards by avoiding the dissemination of misinformation that could have far-reaching consequences.

Beyond privacy and data accuracy, the project navigates additional ethical considerations inherent in conflict mediation research. One such consideration involves the potential for bias in the selection and interpretation of data, especially given the diversity of sources and perspectives in conflict situations. To address this, our methodology incorporates rigorous validation processes and ongoing calibration efforts to mitigate biases and ensure the objectivity of our analysis.



Additionally, we recognize the ethical responsibility to respect cultural nuances and contextual sensitivities when interpreting data. Our commitment to ethical research extends to transparently acknowledging the limitations of our dataset and actively seeking to minimize any inadvertent harm or misrepresentation that may arise from the complexities inherent in conflict mediation research.

Importantly, our data abides by the FAIR data principals. It will be findable and accessible via user-friendly interfaces and data download centre. It can be accessed through PeaceRep and easily through search engines. It will have a detailed codebook making the data itself accessible and a platform which will be accessible for those investigating specific cases or looking for trends. It is interoperable with UCDP (Davies and Engström, 2024), PA-X (Bell et al., 2024), and ACLED (Raleigh et al., 2023) with IDs for conflicts and actors. Mediation events are also geocoded with coordinate information on the event location down to the city level. Finally, it will be reusable and replicable through code and methodology published on the PeaceRep GitHub.

In essence, our ethical considerations extend beyond the automated nature of data collection to encompass the broader ethical landscape of conflict mediation research. By prioritizing privacy, accuracy, and cultural sensitivity, we aim to conduct research that not only meets the highest ethical standards but also contributes responsibly to the understanding of conflict dynamics.

Bias and Fairness in Data Processing

In addressing bias and fairness in data processing, it is essential to recognize the multifaceted nature of biases, both in the automated and manual aspects of our coding process. While we have previously emphasized the dynamic nature of language and translation components and the need for continual refinement, other facets of bias merit attention in the broader context of data processing.



A challenge inherent in our database lies in the exclusion of information from audio and television reporting, limiting our data to written sources. However, in the contemporary information era, we contend that this limitation may not be as severe as it might have been in the past. The prevalence of paper trails and written equivalents to news at a vast scale mitigates many concerns about the exclusion of audio-visual sources. To account for any missed information in this space, we include a vast media transcription and translated database. However, so far, we find no instances of mediation events from audio and visual sources that are not also heavily present on online sources (Internet Archive, 2017).

On the manual side of coding, where human interpretation plays a crucial role, there exists a larger margin for error. Recognizing this, our team is committed to continuous refinement, recoding, and revision of codebook definitions. We employ a double-blind coding methodology, and actively add additional channels to incorporate more detailed and explanatory notes to enhance the interpretive accuracy of our data. This iterative process serves to minimize biases arising from human subjectivity and ensures a more robust and reliable dataset.

Moreover, we acknowledge the potential for biases introduced during the semi-automated coding process. Algorithmic biases, inherent in machine learning models, may inadvertently influence the pulling of data. However, our under-specification of the modelling should account for much of this, and we regularly audit our methods as we build the database.

In our commitment to fairness, transparency, and accuracy, we actively seek to minimize biases at every stage of the data processing pipeline. This includes ongoing calibration sessions to harmonize coding practices among team members and mitigate potential disparities in data.



6 Conclusions & Recommendations

As we conclude this exploration into the semi-automated methods employed to build our new global fragmentation conflict mediation database, we have discussed evolving methodologies, ethical considerations, and our commitment to precision. From the meticulous process of web crawling for conflict mediation references to the nuanced insights derived from our preliminary semi-automated coding, we now conclude with a discussion of the process moving forward, the implications of this work, and some final thoughts.

Future Development and Iterative Review

In envisioning the future development of our conflict mediation research, our commitment to an iterative review process is central to the refinement of our methods and the enhancement of precision and recall in our dataset. The evolving landscape of conflict mediation demands that our methodologies stay attuned to the dynamic nature of geopolitical events and technological innovations, ensuring that our dataset captures the intricacies of ever-changing mediation and informational dynamics.

As technology continues to advance, we anticipate incorporating even more rigorous machine learning models and refining our algorithms. Further, we plan to use our ongoing dataset to continue to train algorithms on our domain. This will not only enhance the efficiency of our data collection and processing but also elevate the accuracy and granularity of our findings. The iterative review process will involve continuous calibration and optimization, aligning our methodologies with the latest advancements in the field, and pulling data that mirrors the emergent trends in conflict mediation. Our commitment to precision and recall extends beyond technological enhancements. We anticipate actively engaging in collaborative discussions within academic and research communities. By presenting our findings at conferences, workshops, and working groups, we invite critical scrutiny and diverse perspectives, fostering a culture of continuous improvement through shared insights and expertise.

Recognizing the inherently dynamic nature of conflict mediation, our team is dedicated to iterative calibration. This involves refining, recoding, and revising codebook definitions, adding new variables, and incorporating nuanced notes to enhance interpretive accuracy. The iterative calibration process ensures that our dataset remains responsive to the evolving landscape of our, and future users' needs.



Implications for Conflict Mediation Research

The implications stemming from our semi-automated conflict mediation database building endeavour are profound, offering a transformative lens through which to understand and engage with the complexities of peacebuilding. These contributions have an impact throughout the broader field of conflict mediation.

First, our meticulous data collection and coding methods allow for a nuanced understanding of the various actors involved in conflict mediation. Beyond recognizing states as primary actors, we identify the roles played by international organizations, non-governmental entities, and individuals. This broader categorization allows for a more comprehensive analysis of the intricate web of actors shaping mediation processes. Further, our work enables a granular examination of mediation scope. We contribute by considering economic and humanitarian dimensions associated with mediation efforts, beyond a mere political scope. Our dataset also sheds light on biases among mediators, offering insights into how various biases may influence mediation events and outcomes.

Finally, considering mediation events as part of broader peace processes is a valuable contribution. This approach helps reveal the coordination and interconnectedness within these efforts. By placing individual mediation events in a larger context, we gain insights into strategic and coordinated actions for sustainable peace. This not only refines our understanding of specific events but also contributes to a more comprehensive view of the complexities involved in peacebuilding. Our research encourages a nuanced exploration that goes beyond isolated events, shaping a more informed conversation on effective conflict resolution strategies.

Closing Thoughts

In conclusion, our report and database development represent a thorough exploration of collective mediation efforts amidst global fragmentation. Leveraging semi-automated coding as a transformative tool, we address the complexities of peace and conflict studies in our ever-evolving world.



We commenced by underscoring the limitations of existing mediation databases and advocating for a more sophisticated approach, especially in the context of emergent global fragmentation. We then delved into our semi-automated methods, covering web crawling, machine learning, ethical considerations, bias in data processing, and an overview of our gathered data.

Our semi-automated methodology is not just a technological advance but a strategic blend of human expertise and machine learning. It tackles conventional challenges faced by traditional databases, enhancing data precision, overcoming language barriers, and adapting to dynamic conflict landscapes.

One area to pursue in this line of research is the recognition and documentation of subnational, alternative, and non-mainstream mediation efforts—an often-neglected aspect in scholarly work. Relatedly, while our dataset primarily captures exogenous conflict-level mediation events, our insights hint at the role of women as informal mediators with rebel groups. Although these instances may not fit into our current dataset, they signify an avenue for future and separate research, urging scholars to explore the informal dynamics of conflict mediation involving women.

The significance of semi-automation lies in providing a more robust, nuanced, and adaptable understanding of mediation processes. As we progress, refining methods and leveraging advancements in machine learning, the implications for conflict mediation research are profound. Our systematic approach marks a significant step forward in navigating the challenges of our complex global landscape.

In essence, the implications of our semi-automated conflict mediation research extend far beyond immediate findings. By offering a comprehensive and nuanced understanding of actors, scopes, biases, and broader peace processes, our research sets the stage for future inquiries that delve deeper into the multifaceted landscape of conflict resolution. Through this work, we aim to inspire a reinvigorated and more informed approach to both the study and practice of conflict mediation.



References

- Belhadj Klaz, S., & Mariani, B. (2022, July 6). Fragmentation of Peacemaking in Syria: Reality and Perception (PeaceRep Report: Global Transitions Series). PeaceRep: The Peace and Conflict Resolution Evidence Platform, University of Edinburgh. <https://peacerep.org/publication/fragmentation-of-peacemaking-in-syria-reality-and-perception/>
- Bell, Christine, Sanja Badanjak, Juline Beaujouan, Tim Epple, Adam Farquhar, Robert Forster, Astrid Jamar, Sean Molloy, Kevin McNicholl, Kathryn Nash, Jan Pospisil, Robert Wilson, and Laura Wise. (2024). PA-X Gender Peace Agreements Database and Dataset, Version 8. www.peaceagreements.org/wggsearch
- Common Crawl. (2008). Overview. In Common Crawl: A web archive for researchers. Retrieved from <https://commoncrawl.org>
- Davies, Shawn, Garoun Engström, Therese Pettersson & Magnus Öberg (2024). Organized violence 1989-2023, and the prevalence of organized crime groups. *Journal of Peace Research* 61(4).
- Duursma, A., & Gamez, S. M. (2023). Introducing the African Peace Processes (APP) dataset: Negotiations and mediation in interstate, intrastate and non-state conflicts in Africa. *Journal of Peace Research*, 60(6), 1010-1020.
- Duursma, A., & Gamez, S. (2024). African Peace Processes (APP) Dataset. Center for Security Studies, ETH Zurich.
- Epple, T., Lang, A., Wise, L., Nash, K., Henry, N., Beaujouan, J., & Desai, D. (2024, July 23). *Navigating peace fragmentation: Recommendations to the new UK government*. PeaceRep. Retrieved from <https://peacerep.org/2024/07/23/navigating-peace-fragmentation/>
- Farquhar, A. (2023, October 31). A Primer on Geocoding for Peace and Conflict Studies. [Peace Analytics Series]. PeaceRep: The Peace and Conflict Resolution Evidence Platform, University of Edinburgh. <https://peacerep.org/publication/primer-on-geocoding/>
- Freer, C. (2022, April 7). MENA Regional Organisations in Peacemaking and Peacebuilding (PeaceRep Global Transitions Series). London: LSE Middle East Centre, The London School of Economics. <https://peacerep.org/publication/mena-regional-organisations-in-peacemaking-and-peacebuilding/>
- Gleditsch, N. P., Wallensteen, P., Eriksson, M., Sollenberg, M., & Strand, H. (2002). Armed conflict 1946-2001: A new dataset. *Journal of Peace Research*, 39(5), 615-637. Uppsala Conflict Data Program UCDP/PRIO Armed Conflict Dataset version 24.1.
- Houghton, K. (2024, October 7). *Conflict and Peacemaking in the Era of Global Fragmentation*. PeaceRep. Retrieved from <https://peacerep.org/2024/10/07/workshop-report-conflict-and-peacemaking-in-the-era-of-global-fragmentation/>
- International Crisis Group (ICG) (2024). CrisisWatch Reports.
- Internet Archive. (2017). Third Eye Data: TV News Archive chyrons. Internet Archive. <https://archive.org/services/third-eye.php>
- Leetaru, K. The GDELT project. GDELT. <https://www.gdeltproject.org/about.html>



- Magara, I. (2023, October 19). *The war in Sudan and the fragilities of regional (non)responses to Africa's armed conflicts*. PeaceRep. Retrieved from <https://peacerep.org/2023/10/19/war-in-sudan-and-fragilities-of-regional-nonresponses/>
- Nash, K. (2021, June 7). *African peace – Regional norms from the OAU to the AU*. PeaceRep. Retrieved from <https://peacerep.org/2021/06/07/african-peace-regional-norms-from-the-oau-to-the-au/>
- Nash, K., Okaro, S., Mugambi Edward, I., & Atambo, L. (2022, March 31). *African regional assistance to South Sudan and Kenya during the pandemic*. PeaceRep. Retrieved from <https://peacerep.org/2022/03/31/african-regional-assistance-to-south-sudan-and-kenya-during-the-covid-19-pandemic/>
- Peter, M. (2024). Global Fragmentation and Collective Security Instruments: Weakening the Liberal International Order From Within. *Politics and Governance*, 12, Article 7357. <https://doi.org/10.17645/pag.7357>
- Raleigh, C., Kishi, R. & Linke, A. (2023). Political instability patterns are obscured by conflict dataset scope conditions, sources, and coding choices. *ACLED Data. Humanit Soc Sci Commun* 10, 74. <https://doi.org/10.1057/s41599-023-01559-4>
- Wise, L. (2023, October 25). *From Protests to Paper: Using PeaceFem to Analyse Women's Inclusion in Peace Processes* [Peace Analytics Series]. PeaceRep: The Peace and Conflict Resolution Evidence Platform, University of Edinburgh. <https://peacerep.org/publication/from-protests-to-paper-using-peacefem-to-analyse-womens-inclusion/>



✉ peacerep@ed.ac.uk

✕ @Peace_Rep_

in PeaceRep

www.peacerep.org
pax.peaceagreements.org

About Us

PeaceRep: The Peace and Conflict Resolution Evidence Platform is a research consortium based at Edinburgh Law School. Our research is rethinking peace and transition processes in the light of changing conflict dynamics, changing demands of inclusion, and changes in patterns of global intervention in conflict and peace/mediation/transition management processes.

Consortium members include: Conciliation Resources, Centre for Trust, Peace and Social Relations (CTPSR) at Coventry University, Edinburgh Law School, International IDEA, LSE Conflict and Civiness Research Group, LSE Middle East Centre, Queens University Belfast, University of St Andrews, University of Stirling, and the World Peace Foundation at Tufts University.

PeaceRep is funded by UK International Development from the UK government.



THE UNIVERSITY
of EDINBURGH



PeaceRep
Peace and Conflict
Resolution Evidence
Platform

PeaceRep: The Peace and Conflict Resolution Evidence Platform
peacerep@ed.ac.uk | <https://peacerep.org> | [@Peace_Rep_](https://twitter.com/Peace_Rep)

University of Edinburgh, School of Law
Old College, South Bridge EH8 9YL