A Primer on Geocoding for Peace and Conflict Studies

Adam Farquhar

PeaceRep
Peace and Conflict Resolution Evidence Platform

PEACE ANALYTICS SERIES

Author: Adam Farquhar

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
@Peace_Rep_
https://www.facebook.com/PeaceRepResearch
https://www.linkedin.com/company/peacerep/

Developed as part of the Peace and Conflict Resolution Evidence Platform, this report intends to support the research programme's work and provide a quick guide to geocoding for the wider field.

This research is supported by the Peace and Conflict Resolution Evidence Platform (PeaceRep), funded by the UK Foreign, Commonwealth & Development Office (FCDO) for the benefit of developing countries. The information and views set out in this publication are those of the authors. Nothing herein constitutes the views of FCDO. Any use of this work should acknowledge the authors and the Peace and Conflict Resolution Evidence Platform.

About the author: Adam Farquhar is a Research Associate and Data Officer at PeaceRep. He supports the management, development, and coding of the PA-X Peace Agreement Database and its sub-databases. His research interests include the application of geocoding and AI in peacebuilding. Adam holds an MA in Comparative Ethnic Conflict from Queens University Belfast and a BA in International Studies from Huntingdon College in Montgomery, Alabama, USA.

The author would like to thank Fiona Knäussel, Christine Bell, and Adam Farquhar for their comments on the report. Thanks to Allyson Doby, and Rick Smith of Smith Design Agency for production work.

Design: Smith Design Agency

Contents

Peace Analytics Series

PeaceRep's Peace Analytics Series features the research methodology underlying the PeaceTech innovations of the PeaceRep programme.

The series includes: data scoping research; 'how to' discussions relating to particular challenges in the field of visualisations and geocoding; and other proof-of-concept tech-based innovations, such as the use of natural language processing. It is intended to present the methodologies and decisions behind our PeaceTech digital research, to make it transparent, and to contribute to establishing a new research digital infrastructure in the field of peace and conflict studies, by supporting others to reuse and repurpose our methodologies and findings.

Introduction

Over the past 30 years, the study of peace and conflict has undergone a 'spatial turn', with a greater emphasis placed on the notion that space is crucial to the overall structure of conflict and peace (Sonja, 1989, Björkdahl and Buckley-Zistel, 2016, p. 3). These spatial characteristics – such as geography, terrain, proximity to centres of power and population density – are pivotal when analysing a variety of research themes related to conflict and peace. Data on politics of peace and conflict are therefore inherently geographic. Tools of geocoding, which associate an event or data point with a latitude and longitude, are increasingly utilized to understand the spatial dimension of peace and conflict.

Geocoded conflict and peace datasets serve to advance this form of research, and combined with qualitative data can enhance our understanding of conflict and peace (Elfvorssen et al., 2020, Macaspac and Moore, 2022). By opening new avenues for investigation, geocoding tools broaden our knowledge of the diverse spatial dimensions inherent in conflict, such as administrative boundaries, political power structures and the intersections of local, national and global politics. These tools also help us to understand the role that technology plays in examining these aspects of conflict that are tied to a specific place or space, and bring to the fore the ethical considerations involved in this new approach to conflict analysis.

The first section of this report explores the concept of geocoding, and how it has been and continues to be used in Politics and International Relations, with a particular focus on Conflict and Peace Studies. Literature in the field is examined and discussed, along with the evolution of geocoding from a primarily geographical tool to a widely used instrument across various fields of study. The section concludes by assessing the benefits and drawbacks of geocoding approaches.

In the second section, the focus turns to a myriad of software tools that can be used in geocoding, as well as the process of taking the geocoded information and turning it into applied geographic data.

From there the third section of the report considers the data available for geocoding, specifically, publicly available data on mapping and administrative boundaries as well as other relevant geocoded datasets.

The fourth section maps out a framework for best practice when geocoding locations and gaining insights from that information. This will be complemented by examples of best practice in data collection in the field, and of ethical issues involved in research using geocoded data.

Finally, the paper provides practical suggestions for working with peace agreements and ceasefires as geocoded locations.

Key Terms

Geocoding refers to taking a description of a location via an address, a coordinate, or a general explanation, and transforming it to a location on the Earth's surface. This results in a geographical feature on a map, with varying attributes. From there the location and attributes can be used for key insights via spatial analysis ("What is Geocoding", n.d.).

What can occasionally lead to confusion, beyond the range of similar terminologies, is the fact that geocoding is sometimes used as a catch-all term for not just the geocoding itself, but the process of mapping, spatial analysis and visualisation. The process can be fluid and separating geocoding from these other tools can sometimes be difficult. As noted by America and Goldberg (2008, p. 6),'...instead of explicitly stating what must be a part of a geocoder, it may be best to leave it open-ended such that different combinations of algorithms and data sources can be employed and still adhere to this definition'. By not being too prescriptive in defining the requirements for geocoding, more room for flexibility and experimentation can allow for increased accuracy and efficiency of the process.

Georeferencing involves taking geographic data and connecting it to a coordinate system to gain insight on that data. An example of this would be taking vector or raster images and placing them on a map of the Earth with the correct scale, often using a geographic information system, or GIS software. While georeferencing could involve using geocoded points, it is the overall process of matching or 'referencing' images or vector data onto a map with the correct location, scale, and geographic coordinate system (Hackeloeer, Klasing, Krisp and Meng, 2014). An example of this in a peace and conflict context would be the UCDP Georeferenced Event Dataset, which connects 'event' phenomena of lethal violence with locations on a map using latitude and longitude coordinates, administrative divisions, and other geographical features such as towns, rivers, or trees, etc. (Högbladh, 2022). Once georeferenced, the spatial dimension of violence can be used to understand its distribution and underlying factors behind the violence.

A **Geographical Information System** (GIS) is a collection of software tools for collecting and analysing geographic data (Gleditsch and Weidmann, 2012). Many of these tools will be further evaluated and discussed in the subsequent sections of this report.

Raster data forms a grid of pixels, each being the smallest unit of a digital image or display, representing a specific location on Earth. Each pixel is assigned a specific value and could denote various geospatial attributes such as temperature, altitude, population density, among others (Esri, n.d.-a). A practical example of raster data in geospatial analysis is a heat map where each pixel corresponds to the density of data points in that location on the map.

Vector data represents a geographic feature using geographic shapes, such as points, lines or polygons. A point could represent a specific location like a country capital; a line could represent a river or a road; a polygon could represent the boundaries around a county or the area in which an event took place. Each vector can contain associated data or attributes that provide additional information about the geographic feature (Esri, n.d.-b).

A **gazetteer** serves two important purposes in the geocoding process. First, in the context of conflict or peace agreement events or when mapping individual locations, a gazetteer helps disambiguate place names or regions. This is particularly useful when a city or town is known by a name in another language or when a location's name or boundary has changed due to a geopolitical event, as is often the case in peace and conflict studies (Geonames.org; Goldberg, Wilson and Knoblock, 2007). Second, it can be used as an initial resource for identifying the latitude and longitude of a location that might already be geocoded.

Finally, **geolocation** has increasingly become part of the everyday lexicon, and can be easily confused with geocoding or georeferencing. Geolocation is the geographic location of an object based on data from an internet-connected device. Types of data source can include mobile phones, cell towers or computer terminals connected to the internet to triangulate a particular point of an object. Geolocation is often used as a tool to help find the geocoded location of a certain point ("What is geolocation", 2020; Yusuf, 2018). An example of this would be to use the location data of a phone which took a picture of an object and determine that object's location from the data. Coverage of conflicts around the world often use location metadata from pictures posted on social media to help determine the location of what is being described in the article.

1 Background

The process of geocoding tracks back to the 1960s, when the United States Census Bureau placed addresses and buildings into different postal codes. This required taking in substantial amounts of data and understanding the spatial relationships between geographic entities. The potential for computer systems to help manage and analyse this geospatial data led to the creation and expansion of Geographic Information Systems (GIS). The '70s and '80s saw these systems refined, leading to their wider adoption by the public from the '90s onward. Since then, the field has expanded to include not just the assignment of coordinates to addresses, but also the geocoding of events, text, and geographical features (Goldberg et al., 2007).

Within the academic literature on peace and conflict studies, the spatial has traditionally been viewed as a less relevant topic of enquiry or belonging to the domain of geographic studies. Massy (1992, p. 66) states that older work 'effectively depoliticise[s] the realm of the spatial'. More recently, however, the spatial element of conflict has been become increasingly important (Goodchild, 2009; Star, 1995; Massey, 1992). In the study of conflict in particular, geographic studies have markedly intensified. This is partly due to advancements in technology, particularly around open-source GIS software, online maps and gazetteers (Gleditsch and Weidmann, 2012, p. 475). In addition, the advent of the internet has made it possible for violent 'events' in question to be easily gleaned from global news reports. One of the more recent technological advancements is the ability to collect and analyse satellite-based data, which has made it possible to use satellite-based night-time light data to analyse various aspects of armed conflicts. This data can be combined with existing datasets to gain insight on the changes in the intensity of conflict over time (Li, Chen and Chen, 2013; Li and Li, 2014; Li, Zhang, Huang and Li, 2015; Li, Li, Xu and Wu, 2017), as well as the causes of inequality among ethnic groups (Bormann, Pengl, Cederman and Weidmann, 2021).

These advancements have led to increasing numbers of geocoded violent events datasets covering most of the world (Elfversson, Gusic, Ha and Meye, 2020, p. 3). While initially focusing on aggregated country-year level analysis, more recent contributions, such as the UCDP Georeferenced Event Dataset, have provided global coverage of disaggregated subnational geocoded violent events (Sundberg & Melander, 2013, p. 524 – 525).

Combining this data with work and methods originating in other disciplines has proven fruitful. For instance, when studying conflict in dense cities, geographic knowledge, along with expertise from the fields of urban studies, criminology, planning, economics, and social anthropology are being incorporated to develop a full understanding of the subject (Elfversson, Gusic and Höglund, 2019).

As with the study of violence and conflict, the area of peace studies has also started to benefit from paying close attention to the geography of events (Björkdahl and Kappler, 2017). Geographers have been able to look at peace and peacebuilding at various levels and scales, such as among different ethnicities or across different covariates which ignore traditional nation state boundaries altogether and focus on other forms of delineation (Björkdahl and Buckley-Zistel, 2016). Examples of such covariates include socioeconomic status, access to resources, population density, and historical or cultural ties between communities. Visualising these factors alongside peace agreements at various scales – ranging from neighbourhoods and cities, to regions, and even broader geographic areas – allows researchers to better understand the geographic distribution of peace and the factors that drive peacebuilding efforts.

The field of international relations has also used spatial information to draw additional insights into geopolitics. The study of the use of public credit in state formation based on location, as well as studies on how conflict is affected by the size and level of engagement of countries that share a border are just a couple of examples (Branch, 2016, p. 849 – 850 Cederman, Gleditsch and Buhaug, 2013). Spatial insight has also enabled challenges to existing ideas around concepts such as civil war. In Inequality, Grievances, and Civil War, by Cederman, Gleditsch, and Buhaug (2013), the authors examine ethnic groups as non-state actors with a spatial position, in between the state and the individual level. This article makes a compelling argument for inequality among ethnic groups as a much likelier influence in the onset of civil war than greed. Geocoding ethnic groups both within and across borders is crucial to help empirically test this argument.

A spatial approach can also help identify alternatives to top-down, elite forms of peacebuilding. In a top-down approach, traditional or colonial ideas of space can hamper a more nuanced and accurate idea of peace, often favouring the liberal state and free markets over the concerns of local communities, whose ideas of peace are steeped in their own culture and history (Autesserre, 2010; Mac Ginty & Richmond, 2013; Richmond, 2009). In Cartographies of Transformation in Mostar and Cape Town: Mapping as a Methodology in Divided Cities by Susan Forde (2019), the post-conflict setting of Mostar, Bosnia and Herzegovina, is given as an example of local residents using derelict spaces as places of engagement across the ethnic divide. An emphasis on this type of engagement can provide an alternative to top-down peacebuilding by showing how local spaces inform peace (Forde, 2020). Indigenous ideas of space and peacebuilding can also contribute to the peacebuilding landscape, particularly where external peacebuilding institutions are present. Indigenous people in Southeast Asia often have an idea of peace that is not reliant on fixed rights based on identity, but on relationships, particularly with the land that they inhabit. A system of gendered spatial relations and local justice can maintain social order when the regional or national state fractures or breaks down (Brigg, George and Higgins, 2022).

Despite its benefits, the use of geocoded data also presents certain limitations and opportunities for improvement. Cities can be a complicated mesh of neighbourhoods and borders that are often contested or understood differently amongst various actors. In these settings, geocoding requires much higher precision (Elfversson et al., 2020). When geocoding conflict, situations can be fluid, and it can be challenging to accurately identify multiple locations of a conflict that could be shifting in and out of the city limits or neighbourhoods. This leads to another complication of simultaneously incorporating time and space into a dataset. The wide use of the polygon information overlay system (PIOS), pioneered in the mid-1970s by the Environmental Systems Research Institute (ESRI), resulted in the prevalence of static spatial information in datasets (Peuquet 2001, p. 13). However, the limitations of static spatial information have led researchers to explore more dynamic approaches to geocoding in conflict data. A notable example is the PRIO-Grid, which will be discussed in more detail later in the report. The PRIO-Grid has the advantage of allowing for daily, weekly, or monthly data collection, aggregated into country-year estimates. This flexibility allows for a more nuanced understanding of the spatiotemporal dynamics of conflict.

Branch (2016, pp. 852 – 865) points to two other issues when using geocoded data, particularly with GIS software: measurement validity and selection bias. Measurement validity can be called into question when the data collected does not accurately portray the political questions being studied. Political institutions and behaviours may not be accurately represented by shape-based maps (vectors) or pixel-based maps (rasters). Shape-based maps represent geographical data with points, lines, and shapes, while pixel-based maps use a grid of small squares (pixels) to show spatial data. An example of this problem can be found in the study of ethnic groups, for instance in the Ethnic Power Relations (EPR) Dataset Family (Vogt et al., 2015). The EPR Datasets provide valuable insights into the relationships between ethnic groups and political power by presenting a comprehensive and systematic analysis of the dynamics governing their political influence. There is heavy emphasis on defining politically relevant ethnic groups and their location in the world. The potential issues around measurement validity that need to be considered are the spatial qualities of ethnic groups themselves. For instance, there is the risk that the dataset might not be used in a way that accurately represent the spatial distribution of ethnic groups, specifically groups that are spread across borders or with large diasporas. Potential concerns also exist regarding the coding and quantification of ethnic groupings, given the inherent complexity and challenges in accurately categorizing such multifaceted identities in a measurable manner (Bochsler et al., 2021; Marquardt, 2021).

These challenges can begin to be addressed in a couple of ways. The latter point can be addressed by using the EPR datasets alongside other datasets to give a more nuanced insight into the nature of the groups being studied. This could include matching with the Varieties of Democracy (V-Dem) dataset, which looks at the underlying inclusiveness of society overall, or matching the groups in the EPR dataset with the Uppsala Conflict Data Program (UCDP) Actor dataset (Marquardt, 2021; Vogt et al., 2015, p. 1338). The spatial distribution issue could be mitigated using three different approaches. First, researchers can replace polygons with points to more accurately represent the location of a group or town, rather than arbitrarily assigning a large space to a group over which it may not have influence. Second, researchers can expand the definition of boundaries. One way of doing this would involve deciding to use boundaries that are jurisdictional (countries, cities, counties etc) or non-jurisdictional (such as areas of fighting or control in civil wars). Selecting the appropriate sub-national unit of analysis by theoretical relevance to the research question can help ensure greater accuracy of results (Soifer, 2019).

Finally, spatial data can be combined with network analysis (Branch, 2016, pp. 853 – 857). This is of particular interest to projects seeking to map networks and combine them with spatial data. One example of this is shown in 'Geographic Determinants of Indiscriminate Violence in Civil Wars' by Sebastian Schutte (2017). In this article, Shutte looks to explore the geographic factors involved in indiscriminate violence within civil wars by using spatial data such as the location of towns, roads, and rivers to understand the spatial distribution of violence, combined with a network analysis of settlements and transportation networks. This approach yields the conclusion that areas with a high density of settlements and poor transportation networks leads to more indiscriminate violent events. This combined method of analysis is promising in peace and conflict studies as it helps to illustrate complex phenomena in a spatial manner.

Selection bias arises when data unrelated to the spatial representation of political institutions is difficult to integrate into analyses, typically through GIS software. This could lead to biased results. Many event-based datasets, such as the Armed Conflict Location & Event Data Project (ACLED), rely heavily on global media reports. Media coverage and accurate location data can vary depending on the remoteness of the location, and both local and external media are susceptible to this inconsistency. The coding of these events can therefore also be affected. Although ACLED supplements its data with expert manual coding, other datasets, such as the Integrated Crisis Early Warning System (ICEWS) and Global Database of Events, Language, and Tone (GDELT), automate event search and inclusion with minimal oversight (Raleigh and Kishi, 2019). It is crucial for researchers to recognize the potential limitations or biases in data collection methods and to consider the evolving nature of media availability and its implications for information selection over time (Miller, Kishi, Raleigh and Dowd, 2022).

Moreover, the representation of political units as polygons on a map may inadvertently cause researchers to neglect significant data that deviates from this format. When data is displayed on a map, the size and coverage of the visual elements might give an impression of comprehensiveness, which could lead to overlooking the impact of groups not spatially represented within states on peace or conflict (Branch 2016, p. 861 – 862).

There are a few proposed solutions to this problem. One involves using different coding methods to combine units that are spatial and those that are not, for instance using not just polygon borders but also points and network analysis to define political units. Another solution involves defining spatial units analytically, rather than by observed real world data. This could involve arranging boundaries into fixed units such as equally sized squares (like squares on a chess board). These equal size squares can make it easier to analyse what is happening in various parts of a country or region where sub-national units can vary a great deal in size and shape, and can also help prevent assumptions about specific areas. The PRIO-Grid is an example of this approach. The PRIO-Grid dataset covers the world in equally sized square polygons. Each polygon is coded with both fixed features, such as terrain, and variable features such as conflict events or population density (Branch 2016, p. 860 – 865).

Geocoding in Peace Studies

Traditionally, there has been a greater emphasis on the study of conflict rather than on the study of peace (Björkdahl and Kappler, 2017, p. 3). This is especially true within geocoding, where there is a growing list of geocoded datasets on conflict, yet few on peace agreements or related topics. While the literature is starting to challenge the dichotomy between conflict and peace, it is true that there are unique challenges when geocoding peace agreements, peace processes, initiatives, and other peace indicators.

One such challenge is the overall utility of geocoding peace agreements outside of the local level. If two countries sign a comprehensive peace deal covering a range of issues, what does pinpointing the spatial element of an inter-state process add to the understanding of these agreements? In a conflict between two or more nation-states, simply pinpointing the countries involved does not improve one's knowledge of peace agreements. If this data was combined with other, more granular data, such as weather patterns, geography (such as contested waters or natural borders) or income level, then more insight could be gained.

Another challenge involves the geographical complexities of local peace agreements. These complexities, across a range of sub-regional to national political settlements, necessitate a nuanced understanding. Geocoding serves as a valuable tool to dissect and comprehend these complexities, thereby enhancing our grasp of the local geographies implicated in such agreements. A pixel or grid-based analysis, for instance, could potentially provide a more precise understanding of the local space and the varied influences of different actors.

This level of detail might help distinguish which local actors have a larger influence that extends to the national or even transnational level, and which actors, particularly in fragmented contexts like Myanmar with thousands of armed groups, have a more limited scope of influence unless they coalesce with a larger group. However, the task of designating an agreement as 'local' is complex, as even conflicts of the smallest magnitude may hold national or transnational implications, and can involve actors whose influence or aspirations extend far beyond the local context (Bell et al., 2021). The spatial interpretation of these local agreements, and specifically the areas affected by them, is crucial to traversing the often-ambiguous boundaries between local, national, and transnational spheres.

Navigating these geographical complexities becomes more achievable when we start to categorize the distinctive types of spaces that local agreements create. According to Bell and Wise (2022), local agreements can create various kinds of real and imagined places. This can include: 'territorially-limited transcalar space', where a defined sub-state area, such as a city, is addressed through local agreements; 'borderland mediation space', which refers to a meeting place between two different groups, like tribes or clans, and where they interact, move, trade, engage in conflict or make peace; and finally 'route-of-passage space', where people not necessarily involved in the conflict, such as displaced populations, nomadic people and aid workers seek to gain passage (p. 567 - 568).

An illustration of local peace geocoding in practice is the PA-X Local dataset (Bell et al., 2023). This dataset compiles written agreements from the principal PA-X Peace Agreements Database that address local issues in some capacity. Part of this dataset includes the latitude and longitude of the local agreement. Instead of representing the geolocation as the place of signing of the agreement, the point on a map represents the epicentre of the conflict addressed in the agreement. If this cannot be determined, the geographic centre of the locale, a central point on a disputed boundary, or the largest local settlement is chosen; in cases where regionally based yet dispersed groups negotiate, the coordinates may be left blank, marked with 999 to denote missing data (Bell et al., 2023, p.13). While this methodology for geocoding has been useful for gaining a better spatial understanding of local conflict, there are some challenges to overcome in representing the complexity of local conflict. For example, a local agreement signed in 2021 between rival communities in South Sudan's Jonglei State involved issues of cattle raiding and child abduction in multiple areas, which is difficult to represent in any granularity with a single point on a map (South Sudan, 2021). Addressing this challenge will be explored further in the report and remains a priority for the PA-X Database going forward (Badanjak, 2021, p. 35).

2 Software Most Commonly Used

Navigating the world of geocoding, with all of its various tools and methods, requires familiarity with many types of software (both free and proprietary), internet applications and programming languages. Exploring this environment for research without earlier experience can be daunting. This section introduces some of the main software, applications and resources needed to tackle spatial research in order, from geocoding data to mapping and analysis.

Two caveats apply to the resources listed below. Firstly, the fields of geocoding and spatial research are constantly evolving with technology and usage. The use of geographic resources in the study of conflict and peace has seen significant uptake in the discipline, which will only increase with time (Cottray, DeYoung, Mills and Upadhyay, 2021).

Resources and technology are constantly changing, with new and improved tools becoming available and often supplanting previous versions. There are a variety of tools that could be used for geocoding, including maps, GIS software and gazetteers, which prevents a one size fits all approach to geocoding (Goldberg et al., 2007).

The second caveat is the issue of open source versus proprietary programmes. Although there are many , open-source resources, others incur a cost – occasionally a substantial one (Gleditsch and Weidmann, 2012) – which is something to be mindful of while considering the sections ahead.

Geocoding

The act of geocoding itself can be performed by an array of programmes that can handle most requirements, from dealing with individual locations or events, all the way to large batch datasets. The complexity of requirements is usually matched by increase in cost, both financial and computational.

■ Google Maps

Google Maps might be the most well-known geocoder. Its free, easy to use API allows the user to enter an address for coding. Checking the terms of service is a key step before use: for example, using the API to geocode a location or event requires using Google Maps to display it (Maps Geocoding API, 2022).

Nominatum

Nominatum is a service that uses Open Street Map (OSM) to geocode locations. The service, which is part of the Open Street Map Project, is open source and intended for occasional use, with bulk geocoding requests discouraged.

QGIS Geocoding Plugins

Quantum GIS (QGIS) is a geographic information system software. There are a few plugins for QGIS that allow you to easily geocode addresses (see Geographic Information System below); these include MMQGIS, OSM Place Search and the Geocoding Plugin. All services are open source licensed.

ArcGIS Online Geocoding Service

This is a paid service which requires credits to use. The addresses to be geocoded need to be compiled into .csv format and imported as a layer on the ArcGIS online service (ArcGIS geocoding documentation). The documentation can be found here: https://doc.arcgis.com/en/arcgis-online/administer/credits.htm.

Map Box Geocoding API

Another useful source for batch geocoding, which is priced per request.

Geocoding with Python

Geocoding can be done using the Python programming language. Packages such as GeoPy and GeoPandas work along with any popular geocoding API such as Google Maps. When using an external service, an API key might be required. You can find the documentation for GeoPy and GeoPandas at https://geopy.readthedocs.io/en/stable/ and https://geopandas.org/en/stable/docs.html.

Geocoding with R

Geocoding can be done with the programming language R using the package “ggmap”. This also uses external services, so obtaining an API key may be required. Documentation can be found at https://www.rdocumentation.org/packages/ggmap/versions/3.0.0.

Edinburgh Geoparser

The Edinburgh Geoparser is a unique programme that can parse through text to geocode locations mentioned therein. It can also generate a timeline of mentioned events. The programme is available at https://programminghistorian.org/en/lessons/geoparsing-text-with-edinburgh.

This list is by no means exhaustive; other services include PositionsStack, TomTom Geocoder and Precisely Geocoding. This list provides several options for getting started with geocoding.

Gazetteers

Geonames.org

A leading gazetteer in academic research. This free and open-source web service provides links between geographic locations and alternate names for those locations. It links this information with attributes such as population, elevation, etc. As this resource is open to public editing, extra care needs to be taken when dealing with less frequently searched locations, as mistakes are less likely to have been noticed by others (Singh, Rafiei, 2018).

Geographic Information System

A Geographic information System (GIS) is a software package that allows the user to import or create geospatial data, link it to a map, analyse the data and manage the analysed data ('What is GIS', n.d.-d). GIS has become increasingly popular in peace and conflict studies and is predominantly utilized via by two main software packages, one proprietary and the other open source.

ArcGIS

ArcGIS is a suite of GIS tools with a wide range of functionalities and is often considered the industry standard (Gleditsch and Weidmann, 2012). While there are lightweight, open-source versions of the software, the main package incurs a cost.

QGIS

QGIS is an open-source GIS tool that is similar to ArcGIS, without the cost and licensing restrictions. Much of the functionality available in ArcGIS is also available in QGIS. A large online community provides support and troubleshooting advice (Santillán, Edwards, Swall and Simmons, 2022).

ArcGIS and QGIS are the only two GIS software packages mentioned specifically in this report as they are the two most likely to be used in academic and policy work related to peace and conflict processes. Other GIS programmes include the geospatial content management system Geonode, as well as SavGIS, GeoDa, SaTScanTM, GWR4 and GAMA (Souris, 2019).

File Formats

Shapefiles

The shapefile format is the most used for geocoding, and the most common format encountered in the field of peace and conflict research. Shapefiles can contain points, lines, or polygon vector data, but only one data type at a time. Points would be used for individual events, buildings, or individuals. Lines are used to demarcate features like rivers or roads. Polygon data is often used to show the boundaries of administrative units such as countries or local administrative borders (Di Salvatore and Ruggeri, 2021, p. 200). However, polygon data can also be used to denote any territory, such as areas under control, areas that have seen fighting, protests, or similar.

Geopackages

A Geopackage is an open format container for geospatial information (Geopackage.org, 2022). It is often used with GIS software such as QGIS. For instance, if one wanted to map the number of listed buildings in Edinburgh, a geopackage format could be used to store the relevant shapefiles, raster files and any other databases or file extensions related to a project.

GeoJSON

GeoJSON is a standard format for displaying geographic features including points, lines and strings. It is based on the JSON format (GeoJSON, n.d.).

Web Map Service

Web Map Service (WMS) files are a protocol for producing georeferenced maps online, developed by the Open Geospatial Consortium. The WMS file format is supported for making maps on QGIS and ArcGIS, as well as other software such as GeoServer, MapServer and Oracle MapViewer (Open Geospatioal Consortium 2022).

Mapping Tools and Map Data

While this report is primarily focused on geocoding, creating online interactive maps of geocoded locations or events can be a valuable tool when deciding how to display your data.

The following is a list of online mapping tools, both paid and free of charge (Open Source):

Open Source

Esri Leaflet

https://esri.github.io/esri-leaflet/
Esri Leaflet is an open source, lightweight set of tools for using ArcGIS services. It uses the Leaflet library, which is a JavaScript library for mobile-friendly interactive maps.

Geojson.io

https://geojson.io/#map=2/20.0/0.0
Geojson.io is an online tool for creating and sharing maps. It uses the GeoJSON format but can use multiple other formats as well. It is particularly useful for drawing polygons around points of interest and exporting them in the Shapefile format.

Open Street Map

https://www.openstreetmap.org/#map=13/7.4032/30.4505
Open Street Map (OSM) is an editable, open-source geographic database of the world. It operates in the same way as Wikipedia in that anyone can contribute, edit or download and export map data. This could include coordinates, information about features represented in a map or other metadata.

- **Open Layers**

https://openlayers.org/

Open Layers is an open-source JavaScript Library for displaying map data in web browsers and mobile applications. It can display vector files, map layers and markers obtained from any source. Open Layers works with multiple formats, including GeoJSON.

Paid/Open Source with Paid Features

- **Mapbox**

https://www.mapbox.com/

Mapbox is a developer-friendly mapping tool that is used for making customized maps and geospatial applications. It has a rich source of spatial data to draw from and is highly customizable.

- **ArcGIS Online**

https://www.arcgis.com/index.html

ArcGIS Online is an online mapping tool that can be used with other ArcGIS products to make maps, analyse data, share, and collaborate. Free accounts are available for non-commercial use.

3 Available Data Resources

While geocoding may seem as straightforward as pinpointing a location on a map, its application in peace and conflict or international relations research contexts can introduce complexity. Such studies often demand a variety of geocoded information at multiple levels, particularly as sub-national research gains popularity in comparative politics, necessitating more granular levels of data (Geroudy et al., 2019; Hallberg, 2012).

Data availability is another important consideration when looking into geocoded information at the sub-national level. The availability of data can have a significant impact on the level of disaggregation that can be reached, as well as on the level at which research can take place within a country (Soifer, 2019, p.108 – 109). If geocoded data on the sub-national administrative units in a country are sparse, then it might make more sense for research to take place at a more general national level. This can be especially true when trying to obtain data outside of the Global North (Lorini, Rando, Saez-Trumper and Castillo, 2020).

The geocoded data required to conduct research could include vector data and the points, lines or polygons that come with it. For instance, analysing instances of violent conflict in a certain region or district of a particular country might require the shapefiles of administrative boundaries, roads or rivers or georeferenced map layers to communicate other valuable information.

The table below provides a list of data resources that cover administrative divisions across different geographical scopes – ranging from global coverage to sub-national units. The resources are categorized based on their area of coverage, such as 'Global,' 'Global Partial,' and 'Sub National,' among others.

Table I: Available global data resources

Name	Area Covered	Notes
GADM Database of Global Administrative Areas	Global	Geopackage, shape file and map data on administrative boundaries around the World
ESRI World Administrative Divisions	Global	Useful map layer providing administrative boundaries
Natural Earth	Global	Free vector and raster map data, with emphasis on natural features
Stack Exchange – Geographic Information Systems	Global	Network supporting GIS work
DIVA-GIS	Global	GIS Software that provides free spatial data including countries and administrative units
Open Street Map Wiki – Potential Data sources	Global	Vast wiki resource on open-source data, contains some out of date information
United Nations – Second Administrative Level Boundaries	Global, Sub National	Official United Nations data on subnational units of counties. Due to technical, political, and practical constraints, not all countries' geospatial data may be available or up to date

Name	Area Covered	Notes
Euro Stat – Geographic Information System of the Commission	Global	Provides up to date spatial data primarily on EU member states, with some data for European countries outside the EU
Xsub	Global/Partial	Cross-references multiple datasets on conflict areas, which could be used to find spatial data in conflict zones
IPUMS International	Global/Partial	Provides shape file administrative units by year (of census)
Rivers as political borders: a new subnational geospatial dataset	Global/Nature	Details instances where rivers make borders on the global, national, and subnational scale
International GIS Data: Global – Penn Libraries – Uni of Pennslyvania	Global	Useful source for global geospatial datasets, maps, and GIS resources
UC San Diego – GIS & Geospatial Technologies: Sorted by Geographic Region	Global	Aggregated source of boundaries and other GIS data, some outdated

Name	Area Covered	Notes
American Uni – Geographic Information Systems & Cartography	Global	Aggregated source of boundaries and other GIS data
OCHA – UDX	Global/Partial	UN aggregated data that include boundaries as well as multiple other datasets
OCHA – UDX Common Operational Dataset Dashboard	*	Useful chart that breaks down data availability of each country
OCHA-UDX Dashboard – How to Guide	*	Help document that details the use of udx data with chosen software
Geoboundaries.org	Global	A large dataset presenting geographical boundaries for the entire world

Notes:
* These resources do not directly include boundaries, but information that states the data's availability or use with different software.

An example of using the above resources to find available data: If a researcher wanted boundary information on the Democratic Republic of Congo, as seen in Figure 1, Geoboundaries.org would be a good place to start. It provides data on the sources of administrative boundaries at levels 0, 1 and 2, organized by licensed permission.

Figure 1:

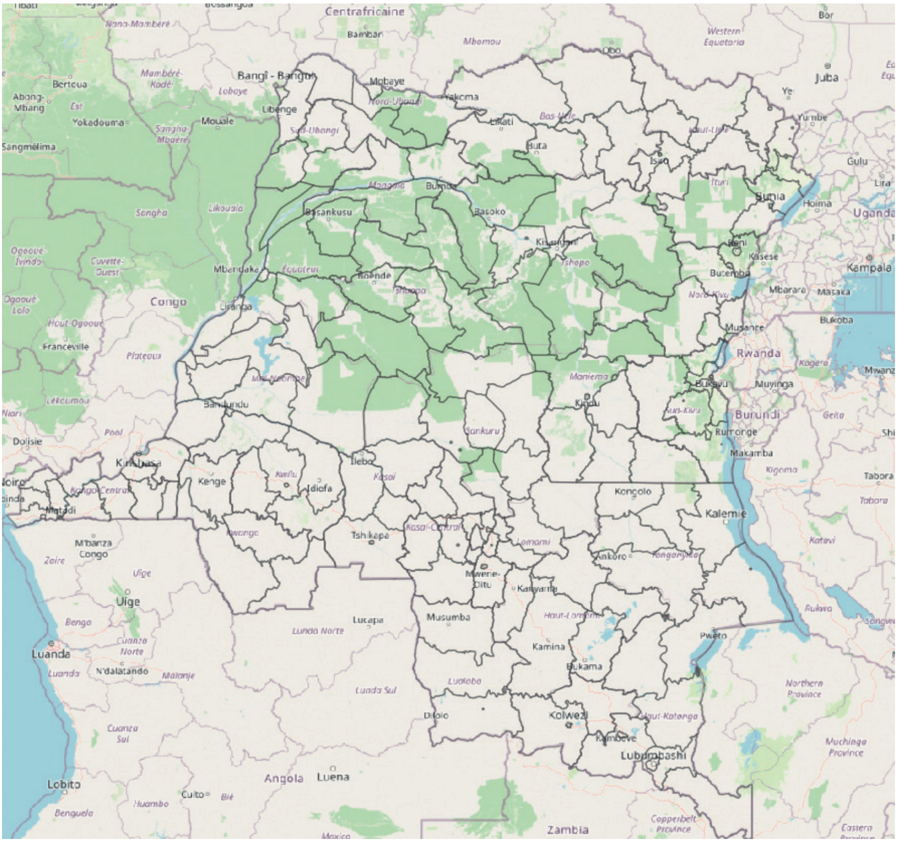

Figure 1: Level 2 Administrative Boundaries of the Democratic Republic of Congo, visualized using the Open Street Map plugin for QGIS, with boundary data sourced from Geoboundaries.org. The illustration enables the integration of other sources of data for comprehensive research. Sources such as Geoboundaries.org, HDX (Provided by the United Nations), and the University of Pennsylvania's Penn Libraries GIS guide are recommended for accurate and updated Shapefile data on administrative units, offering additional geocoded information on global regions.

Geocoded Datasets

The following is a curated selection of datasets that contain geocoded events or instances relating to the field of conflict or peace studies. Collectively, these datasets offer a geographically comprehensive view on the different dimensions of peace and conflict.

Name	Citation	Description
ACLED – Armed Conflict Location and Event Data Project	Raleigh, C., et al. (2010). Introducing ACLED. Journal of Peace Research, 47(5), 651-660.	Provides geocoded information on political violence and protest globally.
Uppsala Conflict Data Program (UCDP)	Davies, Shawn, et al. (2022). Organized violence 1989-2021 and drone warfare. Journal of Peace Research 59(4). Website	Geocoded data on organized violence and civil war worldwide.
Ethnic Power Relations (EPR) Dataset	Vogt, Manuel, et al. (2015). Integrating Data on Ethnicity, Geography, and Conflict. Journal of Conflict Resolution 59(7).	Includes geocoded datasets on ethnic groups and their access to state power.

Name	Citation	Description
Social Conflict Analysis Database (SCAD)	Salehyan, Idean, et al. (2012). Social conflict in Africa: A new database.	Geocoded data on social conflicts in Africa, Mexico, Central America, and the Caribbean from 1990-2017.
Integrated Crisis Early Warning System (ICEWS)	Boschee, Elizabeth, et al. (2015). ICEWS Coded Event Data.	Geocoded database of political events developed by DARPA for early warning of conflict.
Peace Agreement Database (Local)	Bell, Christine, et al. (2020). PA-X Local Peace Agreements Database and Dataset, Version 1.	Part of the PA-X dataset focusing on peace agreements that address local issues, actors, or communal conflict.
Covid-19 Ceasefire Tracker	Allison, J., et al. (2020). An interactive tracker for ceasefires in the time of COVID-19.	Tracks ceasefires that have occurred during the Covid-19 pandemic, including their locations.
The Global Terrorism Database (GTD)	START (2021). Global Terrorism Database (GTD).	Includes geocoded information on over 200,000 terrorist attacks worldwide since 1970.

Name	Citation	Description
Mass Mobilization in Autocracies Database (MMAD)	Weidmann, Nils B., et al. (2019). The Internet and Political Protest in Autocracies.	Contains sub-national datasets on mass mobilization within autocracies around the world.
xSub	Zhukov, Yuri M., et al. (2019). Introducing xSub. Journal of Peace Research 56(4).	Repository of micro-level, subnational event data on armed conflict and political violence from 195 countries between 1968 – 2019.
Global Nonviolent Action Database	N/A. The Global Nonviolent Action Database.	Database of nonviolent action incidents worldwide, with geocoded locations.

4 Geocoding In Practice

This section sets out examples of best practice, focusing on three principal areas:

- Ethics
- Planning and Execution
- Visualisation Considerations and Techniques

Ethics

The ethical concerns of geocoding mirror those of any data collection project, particularly when analysing vulnerable communities or individuals. However, when collecting geographic coordinates of events in conflict or peace-related contexts, there may be additional risks which require specific focus. As in other types of research, ethics in geocoding comes down to 'doing no harm' and trying to prevent harm to anyone associated with the data – although it cannot be assumed that even the best planned project incurs zero risk (Uppsala University, 2021; Solinge et al., 2021). This section will serve as a starting point for a discussion of issues that have recently begun to be addressed with regards to desk research in event-based geospatial data (Green and Cohen, 2020).

The primary concern in collecting geocoded data is making sure no groups or individuals come to harm as a result of data collection or dissemination. The assumption that desk research could be free of many ethical concerns because the researcher is not on the ground collecting data is misguided, especially when it involves event-based datasets that rely on news sources. While the desk researcher may have obtained information in an open and transparent way, the journalists from whom the event data derives may not have done the same. Furthermore, even if a newspaper source is already in the public domain, new patterns or inferences not previously identified in the individual news sources might become apparent when data is aggregated with other event datasets (Green and Cohen, 2020). Moreover, location data can describe events or infrastructure that identify individual people, who have security concerns and a right to privacy, especially among vulnerable groups such as children (Berman, de la Rosa and Accone, 2018, p. 6). This is a crucial consideration as the technology to easily locate and analyse any position on Earth becomes more efficient. Considering potential outcomes, who could be both the victims and perpetrators of harm, and the researcher's capacity for mitigating such danger are all critical to any geocoding risk assessment (van Baalan, 2018).

Another key ethical consideration when geocoding peace or conflict events is whether it brings any real benefit or adds to existing research in a tangible way. Is there something about the geospatial element of what is being researched that could be collected by a less intrusive method, or already available in existing information (Dent et al., 2008, Field, 2022, Berman et al., 2018)? Collecting data for the sake of it may add little value to research while increasing the risk of harm to the subject or area being studied.

Along with the geocoded dataset itself, presentation of the data is another major ethical concern for conducting geospatial research. As Kenneth Field states in his blog post, Ethics and Mapping, "All maps have the power to lie" (Field, 2022, para. 2). Different design choices applied to the same data can change the story that is being told. It is important to consider how visualisations of geocoded data will impact the answers to the research questions being asked.

Finally, increased data collection in academic research has raised concerns about representation. For instance, geocoded datasets that cover areas in the Global South, but which are mainly generated and published at universities in the Global North can exclude the very people who might be the subject of the research, particularly when research concerns countries that have historically experienced colonialism. Under-representation of researchers from the Global South is an ongoing problem that needs to be seriously taken into consideration (Bai, 2018). Some organisations who are currently focused on representation in a geospatial sense include the Spatial Collective, Humanitarian Openstreetmap Team (HOT) and the Worldpop Project.

Planning and Execution

The use of geocoding in peace and conflict studies is new enough that many standards and best practices are still being developed (Cottray et al., 2021). While this offers an opportunity to develop a range of tools and normative practices geared specifically to this area of investigation, it is important to examine best practice from other disciplines or related sectors, in order to understand what has worked in the past and to ensure that mistakes are not replicated going forward. Using other examples can help develop a geocoding plan to guide each step of the design and execution process. This section will provide some guidance in developing a geocoding plan, along with some examples.

It is important to note the difference between a geocoding plan and a data management plan. A geocoding plan, as referenced in this report, lays out the steps taken in converting addresses or events into geographic coordinates. A data management plan sets out how to manage data during and after a project. While the two overlap, the focus of this section is geocoding plans. There are many useful resources available for further information on data management plans (DCC, 2013; Jones, 2011; UK Data Service, n.d.).

Purpose and Feasibility

The first vital step in any geocoding plan is clearly defining its purpose. Why is this tool being used in the first place? Exactly how does this relate back to the research question being asked or the problem to be solved? If geocoding is required for a project, what degree of accuracy is needed (Blossom, 2015)?

One of the downsides to proliferation of technology in the geospatial study of peace and conflict is that it can create the temptation for a research approach that is techno-centric and has a supply-driven methodology (Cottray et al., 2021). When considering the breadth of geocoding, GIS and mapping programmes available, it should always be asked: ‘How does this help answer the research question?’. Often this can be answered without the use of the most up to date, expensive software available – for example, finding the longitude or latitude for the location of a peace settlement can be as straightforward as identifying the coordinates on Google Maps, or by looking them up in a gazetteer.

If a geocoded dataset is required, a feasibility study would be a next step for considering if geocoding produces the required results and accuracy, particularly if working with a large amount of data. This involves taking a subset of the data to be geocoded and testing different methods and software to seek the required results. Publications in the field of public health often include feasability studies and can provide examples and insight into this process (Präger et al., 2019; Baldovin et al., 2015; Pesaresi et al., 2020).

Example: Geocoding Methodology

Aid Data is a research lab at William and Mary's Global Research Institute, focusing on development and foreign aid around the world. As part of that research, Aid Data uses geocoding to track where aid and development come from and to whom they are delivered. The organisation has developed a comprehensive geocoding methodology, outlining the steps involved, including:

- **Reviewing the basic information and sources related to the project or event**
- **Adding additional sources if necessary**
- **Coding specific locations**
- **Reconciling geocoded location from two double-blind coders**

Important points from the methodology include the double-blind approach, the standard used for precision and where to source the coordinates of the locations themselves. Aid Data's double-blind approach means that two researchers are tasked with geocoding a location without any collaboration. The two points are then reconciled by an arbiter (usually a line manager or an automated process). Regarding the level of granularity in coding, Aid Data uses the International Aid Transparency Initiative (IATI) standard, which breaks down the geocoded location to the relevant class level. For example, location class 1 would be an administrative region, while location class 3 would be a structure like a building or bridge. The location is then given a geographic exactness rating of 1 for exact or 2 for approximate. Finally, to source the coordinates themselves, the methodology identifies Geonames gazetteer as the first choice. If the coordinates cannot be found using Geonames, other sources such as Google Maps are suggested.

The methodology provides two further useful recommendations. The first is to be conservative; if coordinates need to be assigned to a larger geographic area in order to capture the phenomenon correctly, this is preferred to assigning an incomplete or arbitrary smaller area or point. The second recommendation is to be as granular as possible by aiming to code meaningful locations rather than centroids of administrative units (Geocoding Methodology, 2017).

Granularity

The next consideration is at what level of granularity to collect your data. This is important for several reasons. Firstly, the level of detail or resolution of the data needs to match the scope of the research question being asked. If there is too much granularity, the results could be affected by 'noise', or arbitrary data that throws off the accuracy of the results. If there is not enough granularity, the results will not provide enough detail to respond to the question (Goas, 2014). Secondly, the data being compared should be at the same scale. If a study is measuring phenomena in different neighbourhoods within a city, various regions within a country, or multiple countries globally, there will often not be the same level of detail available for accurate comparison. It is highly recommended to choose the highest level of granularity that is consistently shared in all areas of the study (Aid Data Research and Evaluation Unit, 2017, p. 12). That way the results will not be skewed by the measurement of different things.

Compatibility

The data must be made compatible with existing datasets. This is important as existing datasets can fill in gaps or be used to confirm the new data's accuracy (Swift and Wilson, 2008, p. 9-10). Even while taking this into account, the research question might require combining data with other data in a different context. Care should be taken in order to account for the context compatibility of the two datasets.

The medical field can provide a good example of this. When investigating how a country's air pollution affects health, a researcher might compare air pollution statistics with hospital statistics (deaths, hospital admissions, etc.). The problem lies in that air pollution data is collected by monitors placed at various locations within the country, therefore measuring pollution around the area of the monitors, but not necessarily spread out evenly across the area being studied. Health statistics, however, are usually aggregated across political boundaries, usually administrative areas. As individual monitoring points and administrative boundaries are spatially mismatched, it should be assumed that either the pollution monitors do measure pollution evenly, or statistical methods need to be employed for accurate comparison of the data (Peng, 2018). This report does not go into detail on these statistical methods, but a good place to start is *Practical Statistics for Data Scientists: 50+ Essential Concepts Using R and Python* by Li-Pang Chen (2021).

Iterative Review

A final consideration is to employ what the Harvard Center for Geographic Analysis call an "iterative mindset" (Blossom, 2015). When geocoding, errors might be found either in the research itself or in the geocoding process. It is important to be prepared to go back – potentially multiple times – to make changes to your geocoding methodology in order to receive more accurate results in accordance with the research aim or question. In fact, an overall scepticism of results can help correct errors and guard against errors turning up again later.

One way of ensuring geocoding consistency is by checking the results against more than one base map (Blossom, 2015). The Uppsala Conflict Database has a three-tier process where data is triple-checked, twice by individuals and once by Python scripts (Sundberg, and Melander, 2013). While highlighting mistakes, however, it also important to highlight what works well, as this good practice will inform research going forward.

Practical Considerations

A number of elements need to be considered to ensure accuracy of results. The first is using the correct coordinate system. There are two types of coordinate systems: geographic and projected. Geographic coordinate systems relate to locations on a 3D model of Earth and are measured in decimal degrees. A projected coordinate system refers to a 2D model of Earth, usually in the form of a map, which is measured in linear units. There are different subtypes within both systems, and differing indirect ways of converting data from one system to another, which means attention needs to be paid to which geographic system a potential data source relates to. When working with spatial data, coordinate systems often vary between sources. It is crucial to account for this variation to guard against inaccuracy of results (Di Salvatore and Ruggeri, 2021; Smith, 2020).

Another aspect to consider is the potential for temporal bias. Access to spatial data has improved significantly over time, particularly in regions of the Global South, meaning that the accuracy of the data may diminish the further back in time. One way to identify temporal bias is by looking at administrative subdivisions of the countries or regions of interest. A significant increase in the amount of data available could indicate temporal bias, as the information available may have increased over time (Rosvold and Buhaug, 2021, p. 3).

Visualisation Considerations and Techniques

Visualising geocoded data on complex processes such as peacebuilding can be challenging. There is no 'one size fits all' approach to addressing this level of complexity, and there are a variety of tools that can be used in different scenarios. The following are some considerations and techniques to keep in mind when approaching visualisations.

Open Source or Paid Tools?

Often lightweight, meaning user-friendly, and with low memory consumption, mapping software (such as Google Maps or Nominatum) is sufficient for mapping geocoded data. However, to map a large volume of data or to undertake a complex project, GIS software might be necessary. When choosing which GIS software to use, you will be likely be faced with two choices: ESRI ArcGIS or QGIS. ArcGIS, as described earlier in this report, is the industry standard when it comes to spatial analysis because of its breadth and scope. One of the advantages of using ArcGIS online software is that the user can build interactive maps that can then be exported or embedded as a dashboard on a webpage. One notable example of this is John Hopkins University COVID-19 Dashboard and Global Map. The license is available at a lower cost for non-profit organisations (AFP YouTube Wkshp, 2021). The dashboard also provides links to open-source tools such as the GIS 4 Peace hub, a self-service platform that provides access to open data, mapping tools and case studies to support and better inform peacebuilding initiatives (ArcGIS, n.d.). The downside to the main ArcGIS platform is that its market-leading position is reflected in its price; the high cost of the platform could prove prohibitive to users. On top of this, most of the training in GIS provided by Esri is related specifically to its software (Gleditsch and Weidmann, 2012).

ArcGIS also provides a tool called StoryMaps, which combines text with map visualisations to create polished, interactive content (Esri, n.b.-c). When visualising complex processes, more contextual information may be required. Displaying geocoded maps along with text, video and other types of graphs in a visually engaging way can communicate the nuance that is often present in spatial data. It should be noted that StoryMaps is part of the propriety package of the ESRI ArcGIS software.

The open-source alternative to ArcGIS is QGIS, which has developed to rival many capabilities of the proprietary ArcGIS package. Most importantly there is a large amount of support available online for most tasks using QGIS, as well as support for the many plugins that are available for free (Gleditsch and Weidmann, 2012; Santillán et al., 2022).

Most research tasks involving use of GIS by an individual researcher – and many by organisations – can be achieved with the open-source GIS option, saving money and time for researchers to familiarise themselves with the software.

Visualising Boundaries

Often the boundaries of what is being geocoded are not adequately represented by individual points, but rather by events, roads and infrastructure, or natural features such as rivers and lakes. A potential solution would be to use the Geojson.io tool discussed in the 'Mapping tools and map data' section, which enables the user to draw enclosed shapes or lines and allows coordinates of these shapes to be exported in multiple formats, including GeoJSON, CSV and Shapefile formats. The learning curve for this web-based tool is small, it does not incur a cost, and is perfectly suitable for small to medium sized geocoding tasks. Another type of visualisation of different vector and raster data is the PRIO-GRID, developed by Andreas Forø Tollefsen, Håvard Strand and Halvard Buhaug at the Peace Research Institute in Oslo. It is an example of best practice in displaying disaggregated information at a more localised level.

The PRIO-GRID provides a global spatial grid structure at a measurement of 0.5 x 0.5 decimal degrees, in order to display data both static and temporal in nature (Di Salvatore and Ruggeri, 2021; Tollefsen, Strand and Buhaug, 2012; Nemeth, Mauslein and Stapley, 2014; Rosvold and Buhaug, 2021). As illustrated in Figure 2, the grid structure can display multiple datasets from various sources, with categories of data ranging from wealth distribution, weather patterns, elevation, population density and conflict, to name a few.

There are many advantages to this type of approach. Because it is free of national boundaries, it is free of the biases that these boundaries can cause. This is especially true when working with events that either take place locally within a country's boundaries, or transnationally. With the use of the 0.5 x 0.5 grid cells, the data can be scaled up or down accordingly.

An example of this would involve analysing peace agreements or ceasefires. While it is true some peace agreements that take place purely at a national level and can simply be assigned to the country's capital or centroid, many are also regional, local or transnational in nature (Bell et al., 2021; Wise et al., 2021). Furthermore, even if the agreement is coded to one country, it could still have a strong connection to an event or agreement coded to another country. As stated above, any practice of placing a single point vector as a geolocated location of a peace agreement would fail to lend itself to the sort of analysis required for studying these agreements locally and transnationally. Such analysis could also be combined with other datasets which use the PRIO-GRID, such as the Uppsala Conflict Data Programme (UCDP).

Perhaps the most ambitious potential benefit of the PRIO-GRID dataset is its aim to standardize spatial analysis in the field of conflict studies, although the authors of this system acknowledge that it would most likely complement other types of datasets depending on the nature of the research itself. Some further development on the grid has involved including software that allows users to easily build the grid for research, as well as consideration of the size of the grid cells and how that can affect analysis (Pickering, 2016; Suzuki, 2022).

Figure 2:

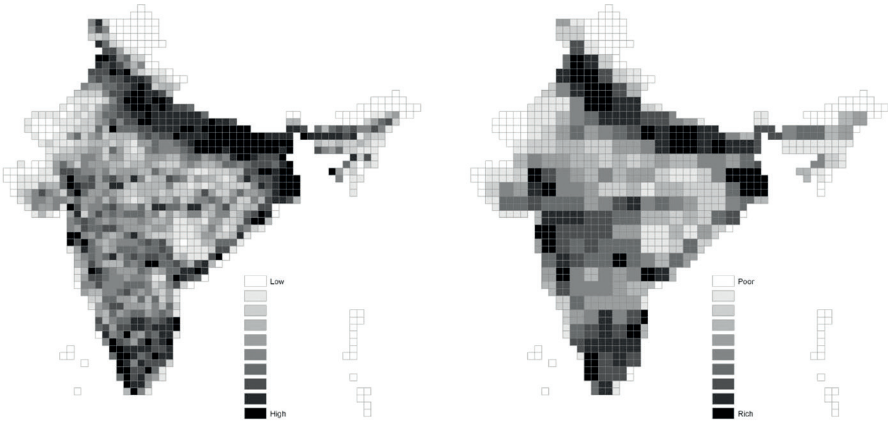

Figure 2: PRIO-GRID, adapted from "PRIO-GRID: A Unified Spatial Data Structure" by A. F. Tollesen, Journal of Peace Research, 2012. This figure shows the local population density (left) and the spatial distribution of wealth (right) for India in 1990. Areas with darker shading signify regions with a greater population and increased income, respectively. The PRIO-GRID database is recommended for high-resolution geocoded data. Source: Journal of Peace Research, 2012, 49(2): 363-374. Used under Creative Commons Attribution-Non Commercial 3.0 License.

Best Practice Examples

This section discusses examples of studies and projects that have employed geocoding best practice in their work.

Example 1: Using GPS Data to Geocode Movement

As mentioned in the definition of geocoding at the beginning of this report, there are other ways to geocode beyond using online maps. Global Positioning System (GPS), which some consider the gold standard in geocoding, can be used effectively to gain insight into peace and conflict studies (Goldberg et al., 2007). Specifically, GPS can help us understand the movements and interactions of people or communities in areas with ongoing disputes or tension. This is important because how people view and navigate these 'contested spaces' can tell us a lot about the underlying conflicts. Sometimes, what people understand as interfaces in these areas can differ significantly from the official, mapped interfaces. GPS can help us capture these nuances.

An example of this is Raanan and Shoval's (2014) examination of spatial activity and perceived boundaries in Mental maps compared to actual spatial behaviour using GPS data: A new method for investigating segregation in cities. The authors were interested in how people's perceptions of contested space affect their everyday movements, and how these movements and perceptions stacked up against assumptions of division within the city of Jerusalem.

The authors highlight three areas touched on by this report, making it a good example of best practice: the preparation for geocoding, georeferencing using traditional methods of data collection to compliment geocoding methods, and the use of GPS to geocode movement in a divided society.

Preparatory work for the geocoding involved interviews over a five-month period with all 18 participants to discuss their perceptions of boundaries and space in their everyday lives. Participants drew on blank maps of the city of Jerusalem to locate areas they were familiar with, identifying key neighbourhoods and areas where they did and did not feel comfortable. After the interview, each participant was given a GPS tracker for a week. This is a good example of some of the preparatory work that must be factored into a geocoding methodology, especially when geocoding involves work on the ground with individuals.

The next part of this study highlighted georeferencing, which can sometimes be confused with geocoding. The maps that were drawn on by interview participants were scanned, and then georeferenced using ArcInfo 9.3 software, which was part of the ArcGIS desktop software package (see Geographic Information Systems). The participants' movement data, as recorded by the GPS tracker, were made into polygons and added to the maps as new layers to be analysed. This process was undertaken as soon after the interviews as possible, so that maps that were annotated in the interviews would match the participants' comments as accurately as possible. The maps were then sent to each participant to confirm that the final map did not include any misinterpretations.

The use of the GPS tracking is a good example of combining the technical process of geocoding with more traditional methods of data collection in the social sciences. The tracking data itself was divided into tracks, (paths the participants took to and from places like work and home), and activity nodes (places where the participant spent longer than fifteen minutes). The tracks and activities nodes were then added to the georeferenced map after the information had been cross-referenced with activity diaries. This turned out to be important, as it flagged up issues such as bus stop waits that were well over 15 minutes but did not constitute an activity, or doctor visits that took less than 15 minutes and therefore did not register as an activity.

A project of this nature inherently brings forth a range of security considerations that warrant careful attention and consideration. It would be a challenge to conduct this type of research in a safe manner in most locations affected by armed conflict. This best practice example shows, however, that the process of geocoding – especially when working in such an intimate context – should be combined with as much on the ground expertise and information as possible. It is also important that geocoded data be supported by other metadata; in this case, the activity journals and interviews with participants.

Example 2: Polygons and Grids for Clusters of Conflict

Buhaug and Rød's paper entitled Local determinants of African civil wars, 1970-2001 (2006) looks at conflict data from Uppsala/PRIO Armed Conflict Dataset to identify clusters of conflict that correlate with spatial distribution of influencing factors. This research highlights the use of spatial research at a sub-national level, using drawn polygons as well as grids as units of analysis. This is useful because the supporting geo-referenced data are as easily converted to grid cells as they are to administrative units. Additionally, instead of creating a single point of conflict around which a radius is drawn, the authors use GIS software to draw polygons indicating the area of the specific conflict zone. This subnational georeferenced data enabled the researchers to connect intrastate conflicts to factors such as population density, terrain and industry.

Part of what makes this example useful for geocoding peace agreements or ceasefires is that one could quite easily use the Geojson.io tool (see 'Map tools and map data' section) to draw polygons around areas related to peace agreements, and then export them as Shapefiles. From there, this data could be considered along with other exogenous factors for analysis.

Regional Data from Global Datasets

This report has focused on sub-national geocoded data. But what if one wanted to work with regional/interstate geocoded data? One potential challenge is the difficulty obtaining relevant disaggregated data for the regions of interest. Most studies on this scale draw from global datasets (De Juan, 2012, p. 12).

An example of drawing regional figures from global datasets is Raleigh and Kniveton's work entitled *Come rain or shine: An analysis of conflict and climate variability in East Africa* (2012). In this paper the authors investigate the relationship between conflict and climate, responding to competing claims about how much the former is influenced by the latter. The paper refers to the geocoded locations of conflicts in the three East African countries of Uganda, Kenya and Ethiopia, using geocoded data from the Armed Conflict Location and Event dataset (ACLED). This study looked at monthly data points over a 13-year period in order to understand how the number of conflicts varied through different seasons of rainfall over time. This East African region was chosen due to the availability of data on conflict in the area, as well as the consensus around the impacts of climate change on the region.

The paper ultimately reported that, in this particular case, there appeared to be a link between rainfall variation and an increase in different types of conflict, depending on whether there was a significant lack of rainfall or an increase. This highlights the importance of looking at a cross-national geocoded datasets of local conflict events, in order to inform debates that rely heavily on large scale conflicts. For further examples of articles that use geocoded datasets, see Appendix.

Conclusion and Recommendations

Technical Recommendations

Tracking method of geocoding

When coding a peace agreement or ceasefire, a clear hierarchy of geocoding decisions should be followed for consistency. At the point at which the hierarchy is defined, there should be a box to tick so that this data can be easily traced and recognized by end users. Ultimately, this approach will enhance the consistency and clarity of geocoding decision-making.

Double-blind geocoding of locations for accuracy

Double-blind geocoding should take place, with a third individual reconciling the two.

The use of PRIO-GRID style maps to visualise local or cross-national peace agreements and ceasefires

The PRIO-GRID style data discussed in this report can be produced with programmes such as SpatialGridBuilder. QGIS software or R packages can also be used. These tools should be explored in a trial run when seeking to visualise multiple local peace agreements that span different administrative units or countries.

Use geojson.io for polygon shapes to indicate areas not easily marked by a single point.

Geojson.io is a simple programme for drawing polygon shapes around areas covered by peace agreements or ceasefires. This would be especially useful for local agreements that cover more than one location. on the limitations of drawing boundaries around these areas would need to be made clear, as they may not represent administrative boundaries of the areas being covered. This could be mitigated by providing further non-spatial metadata as additional context for drawn areas.

Programme Recommendations for PeaceRep and Peace/ Conflict Researchers

■ Geocoding peace agreements should remain at the local level or for individual country studies

Peace agreements at the national level may not have much to gain from geocoding, unless something specific about local geographies between or within the countries involved can be accurately captured with geocoding. As such, geocoding the complete PA-X database is not necessary. Local peace agreements, individual country reports and publications are examples where the academic value of geocoding approaches could be useful.

■ Basic training in the use of GIS for research going forward

As the spatial dynamics of peace and conflict are becoming more important, and the technology behind this research focus is constantly improving, it would be useful to equip members of the PeaceRep team with a foundational knowledge of GIS and how it could be used for our work.

■ Incorporate the use of ArcGIS StoryMaps in PeaceRep reporting

This tool can be a useful way to communicate complex geospatial dimensions of the peace agreements and conflicts we look at, especially those of a local nature.

■ Annual review of uses and methods of geocoding as the field and technology progresses

Geocoding in this field will be progressing constantly. Reviewing current consensus on best practice and methods, as well as available technologies, will be required on an annual basis.

Conclusion

The ‘spatial turn’ in the study of peace and conflict has led to a growing number of researchers harnessing geocoding tools to better understand these disciplines. The study of space in the political sphere is no longer limited to the geographer; geocoding is a skill for anyone seeking to better understand how conflicts are fought and how peace is won. This opportunity brings a need for a full understanding of the promises and limits of such a tool, as well as the impact it can have on the communities being studied.

This report summarised existing literature on geocoding and the potential for insight that the spatial dimension can provide researchers in this field, while highlighting a need for awareness around mischaracterisation of events being studied through measurement validity or selection bias. An overview of the software, programmes, and data available for geocoding highlighted the growing number of tools available to the researcher. While proprietary software and programmes are available to address geocoding needs, options that are open-source and free of charge would not only suffice in many cases, but might be preferred. Examples of research studies that have employed geocoding methodology in various ways were shared, as well as considerations for best practice, particularly around ethical concerns around the sensitive nature of the data and potential effects on people’s lives; identifying the latitude and longitude for a location or event might be a desk-based exercise, but real-life consequences and ethical quandaries may be involved. Recommendations were provided for the PeaceRep programme on making the best use of geocoding tools going forward, including recommendations for use of the geojson.io application for creating polygons to represent peace events on a map, and for considering the use of the PRIO-GRID to effectively illustrate local or transnational agreements across administrative lines. This report also recommends that the PeaceRep programme continues to focus on geocoding local peace agreements going forward, as efforts to geocode interstate agreements would likely raise practical issues that could nullify the significant insights that geocoding can provide.

References

Administration Level Definitions. (n.d.). Socioeconomic Data and Applications Center (sedac).Retrieved from: https://sedac.ciesin.columbia.edu/povmap/ds_defs_admin.jsp

Aid Data Research and Evaluation Unit. (2017). Geocoding Methodology (Version 2.0.2). Aid Data. http://docs.aiddata.org/ad4/files/geocoding-methodology-updated-2017-06.pdf

Allison, J., Badanjak, S., Bach, B., Bell, C., Bhattacharya, D., Knaussel, F., & Wise, L. (2020). An interactive tracker for ceasefires in the time of COVID-19. The Lancet Infectious Diseases. https://doi.org/10.1016/S1473-3099(20)30932-4

America, N., & Goldberg, D. (2008). A Geocoding Best Practices Guide. University of Southern California GIS Laboratory. Submitted to North American Association of Central Cancer Registries, Inc.

ArcGIS. (n.d.). GIS for Peace - DSSD Global. Retrieved May 25, 2023, from https://gis-4-peace-dssdglobal.hub.arcgis.com/

Autesserre, S. (2010). The trouble with the Congo: Local violence and the failure of international peacebuilding. Cambridge University Press.

Badanjak, S. (2021). The PA-X Peace Agreement Database: Reflections on documenting the practice of peacemaking. Pathways to Peace and Security, 2(61). pp. 24-42. https://doi.org/10.20542/2307-1494-2021-2-24-42

Bai, Y. (2018). Has the Global South become a playground for Western scholars in information and communication technologies for development? Evidence from a three-journal analysis. Scientometrics, 116(3), 2139–2153. https://doi.org/10.1007/s11192-018-2839-y.

Baldovin, T., Zangrando, D., Casale, P., Ferrarese, F., Bertoncello, C., Buja, A., Marcolongo, A., & Baldo, V. (2015). Geocoding health data with geographic information systems: A pilot study in northeast Italy for developing a standardized data-acquiring format. Journal of Preventive Medicine and Hygiene, 56(2), E88–E94. https://doi.org/10.15167/2421-4248/jpmh2015.56.2.442

Bell, C., Badanjak, S., Beaujouan, J., Epple, T., Forster, R., Jamar, A., Molloy, S., McNicholl, K., Nash, K., Pospisil, J., Wilson, R., & Wise, L. (2021). PA-X Peace Agreements Database and Dataset, Version 5. www.peaceagreements.org

Bell, C., & Wise, L. (2022). The Spaces of Local Agreements: Towards a New Imaginary of the Peace Process. Journal of Intervention and Statebuilding, 16(5), 563-583. https://doi.org/10.1080/17502977.2022.2156111

Bell, C., Badanjak, S., Wise, L., Adhikari, M., Beaujouan, J., Distrotti, M., Epple, T., Forster, R., & Wilson, R. (2023). PA-X Local Peace Agreements Database and Dataset, Version 4. www.peaceagreements.org/lsearch

Berman, G., de la Rosa, S., & Accone, T. (2018). Ethical Considerations When Using Geospatial Technologies for Evidence Generation. Innocenti Discussion Papers no. 2018-02.

Björkdahl, A., & Buckley-Zistel, S. (2016). Spatializing Peace and Conflict: An Introduction. In *Spatializing Peace and Conflict* (pp. 1–22). Palgrave Macmillan UK. https://doi.org/10.1057/9781137550484_1

Björkdahl, A., & Kappler, S. (2017). Peacebuilding and spatial transformation: peace, space and place. Routledge. https://doi.org/10.4324/9781315684529

Blossom, Jeff. (2015). Geocoding Best Practices. Center for Geographic Analysis Harvard University. https://gis2.harvard.edu/services/blog/geocoding-best-practices

Buhaug, H. & Rød, J. K. (2006). Local determinants of African civil wars, 1970–2001. *Political Geography, 25*(3), 315–335. https://doi.org/10.1016/j.polgeo.2006.02.005

Bormann, N., Pengl, Y., Cederman, L., & Weidmann, N. (2021). Globalization, Institutions, and Ethnic Inequality. International Organization, 75(3), 665-697. doi:10.1017/S0020818321000096

Boschee, E., Lautenschlager, J., O'Brien, S., Shellman, S., Starz, J., & Ward, M. (2015). "ICEWS Coded Event Data". Harvard Dataverse, V37, UNF:6:NOSHB7wyt0SQ8sMg7+w38w== [fileUNF]. https://doi.org/10.7910/DVN/28075

Bochsler, Green, E., Jenne, E., Mylonas, H., & Wimmer, A. (2021). Exchange on the quantitative measurement of ethnic and national identity. Nations and Nationalism, 27(1), 22–40. https://doi.org/10.1111/nana.12688

Branch, J. (2016). Geographic Information Systems (GIS) in International Relations. International Organization, 70(4), 845–869. http://www.jstor.org/stable/44651925

Brigg, M., George, N., & Higgins, K. (2022). "Making Space for Indigenous Approaches in the Southwest Pacific? The Spatial Politics of Peace Scholarship and Practice." Journal of Intervention and Statebuilding 16 (5): 545–562. doi:10.1080/17502977.2022.2139908

Cederman, Gleditsch, K. S., & Buhaug, H. (Eds). (2013). Inequality, grievances, and civil war. Cambridge University Press.

Chen, L-P. (2021). Practical statistics for data scientists: 50+ Essential concepts using R and Python. Technometrics, 63(2), 272-273. https://doi.org/10.1080/00401706.2021.1904738

Cottray O., DeYoung S., Mills L., Upadhyay S. (2021, December 1-2). Advancing the use of GIS in Peacebuilding [Conference workshop]. Emerging Technologies in Peacebuilding and Prevention Virtual Practitioners Workshop, New York, NY, United States. https://www.youtube.com/watch?v=zhCXAC9Ma1o

DCC. (2013). Checklist for a Data Management Plan. v.4.0. Edinburgh: Digital Curation Centre. Available online: http://www.dcc.ac.uk/resources/data-management-plans

Dent, B., Torgusen, J., & Hodler, T.W. (2008). Cartography: Thematic Map Design. McGraw-Hill.

De Juan, A. (2012). Mapping Political Violence – The Approaches and Conceptual Challenges of Subnational Geospatial Analyses of Intrastate Conflict. German Institute of Global and Area Studies (GIGA). http://www.jstor.org/stable/resrep07595

Di Salvatore, J., & Ruggeri, A. (2021). Spatial analysis for political scientists. Italian Political Science Review/Rivista Italiana Di Scienza Politica, 51(2), 198–214. https://doi.org/DOI: 10.1017/ipo.2021.7

Dorsey, S. (2020). The opportunity cost of intrastate violence and the out-of-sample validity of commodity price shocks. The Journal of Defense Modeling and Simulation, 17(3), 309–324. https://doi.org/10.1177/1548512919837383

Dong, E., Du, H., & Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. The Lancet infectious diseases, 20(5), 533-534.

Elfversson, E., Gusic, I., & Höglund, K. (2019). The spatiality of violence in post-war cities. Third World Thematics : a TWQ Journal, 4(2-3), 81–93. https://doi.org/10.1080/23802014.2019.1675533

Elfversson, E., Gusic, I., Ha, T.-N., & Meye, M.-T. (2020). Geocoding as a Method for Mapping Conflict-Related Violence. In The Palgrave Encyclopaedia of Peace and Conflict Studies (pp. 1–11). Springer International Publishing. https://doi.org/10.1007/978-3-030-11795-5_214-1

Esri. (n.d.-a). What is raster data? ArcGIS Desktop. Retrieved May 29, 2023, from https://desktop.arcgis.com/en/arcmap/latest/manage-data/raster-and-images/what-is-raster-data.htm

Esri. (n.d.-b). What is vector data? ArcGIS Desktop. Retrieved May 29, 2023, from https://desktop.arcgis.com/en/arcmap/latest/manage-data/overview/what-is-vector-data.htm

Esri. (n.d.-c). ArcGIS StoryMaps. Retrieved June 14, 2023, from https://www.esri.com/en-us/arcgis/products/arcgis-storymaps/overview

Esri. (n.d.-d). What is GIS? Retrieved February 21, 2022, from https://www.esri.com/en-us/what-is-gis/overview

Field, K. (2022). Ethics in Mapping. ArcGIS Blog. https://www.esri.com/arcgis-blog/products/arcgis-pro/mapping/ethics-in-mapping/

Forde, S. (2019). Cartographies of Transformation in Mostar and Cape Town: Mapping as a Methodology in Divided Cities. Journal of Intervention and Statebuilding, 13(2), 139–157. https://doi.org/10.1080/17502977.2018.1528713

Forde, S. (2020). Human Geography and Peacebuilding. In The Palgrave Encyclopaedia of Peace and Conflict Studies (pp. 1–10). Springer International Publishing. https://doi.org/10.1007/978-3-030-11795-5_52-1

GeoJSON. (n.d.). Retrieved from http://www.geojson.org

Giraudy, Moncada, E., & Snyder, R. (2019). Inside Countries: Subnational Research in Comparative Politics. Cambridge University Press. https://doi.org/10.1017/9781108678384

Gizelis, T. I., & Cao, X. (2020). A security dividend: Peacekeeping and maternal health outcomes and access: https://Doi-Org.Ezproxy.Is.Ed.Ac.Uk/10.1177/0022343320917198, 58(2), 263–278. https://doi.org/10.1177/0022343320917198

Gleditsch, K. S., & Weidmann, N. B. (2012). Richardson in the Information Age: Geographic Information Systems and Spatial Data in International Studies. https://doi.org/10.1146/annurev-polisci-031710-112604

Goas, V. (2014, August 19). Geocoding 101: A Behind the Scenes Look at Visualizing Aid and Development. The First Tranche. https://www.aiddata.org/blog/geocoding-101-a-behind-the-scenes-look-at-visualizing-aid-and-development.

Goldberg, Wilson, J. P., & Knoblock, C. A. (2007). From text to geographic coordinates: The current state of geocoding. Journal of the Urban and Regional Information Systems Association, 19(1), 33–46.

Google Developers. 2022. Maps Geocoding API. [online] Available at: https://developers.google.com/maps/documentation/geocoding/overview#Geocoding [Accessed 2 February 2022]

Gravitate Team. (2020). What is Geolocation? Blog: https://www.gravitatedesign.com/blog/what-is-geolocation/

Green, H., Cohen, D. K. (2021). Centering Human Subjects: The Ethics of "Desk Research" on Political Violence. Journal of Global Security Studies, 6(2). https://doi.org/10.1093/jogss/ogaa029

Harbers, Ingram, M., Giraudy, A., Moncada, E., & Snyder, R. (2019). Politics in Space: Methodological Considerations for Taking Space Seriously in Subnational Research. In Inside Countries (pp. 57–91). Cambridge University Press.

Hackeloeer, A. Klasing, K., Krisp, J. M., & Meng, L. (2014). Georeferencing: a review of methods and applications. Annals of GIS, 20(1), 61–69. https://doi.org/10.1080/19475683.2013.868826

Hallberg, D. J. (2012). PRIO Conflict Site 1989–2008: A Geo-Referenced Dataset on Armed Conflict. Conflict Management and Peace Science, 29(2), 219–232. https://doi.org/10.1177/0738894211433168.

Jones, S. (2011). 'How to Develop a Data Management and Sharing Plan'. DCC How-to Guides. Edinburgh: Digital Curation Centre. Available online: /resources/how-guides

Levin, J. (Ed.). (2020). *Nomad-State Relationships in International Relations Before and After Borders.* Springer International Publishing. https://doi.org/10.1007/978-3-030-28053-6

Li, X., Chen, F., & Chen, X. (2013). Satellite-Observed Nighttime Light Variation as Evidence for Global Armed Conflicts. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 6. 2302-2315. 10.1109/JSTARS.2013.2241021.

Li, X., & Li, D. (2014). Can night-time light images play a role in evaluating the Syrian Crisis? *International Journal of Remote Sensing, 35*(18), 6648–6661. https://doi.org/10.1080/01431161.2014.971469

Li, X., Li, D., Xu, H., & Wu, C. (2017). Intercalibration between DMSP/OLS and VIIRS night-time light images to evaluate city light dynamics of Syria's major human settlement during Syrian Civil War. *International Journal of Remote Sensing, 38*(21), 5934–5951. https://doi.org/10.1080/01431161.2017.1331476

Li, X., Zhang, R., Huang, C., & Li, D. (2015). Detecting 2014 Northern Iraq Insurgency using night-time light imagery. *International Journal of Remote Sensing, 36*(13), 3446-3458. https://doi.org/10.1080/01431161.2015.1059968

Lorini, V., Rando, J., Saez-Trumper, D., & Castillo, C. (2020). Uneven Coverage of Natural Disasters in Wikipedia: The Case of Flood. arXiv:2001.08810v1 [cs.IR]. https://doi.org/10.48550/arXiv.2001.08810

Macaspac, & Moore, A. (2022). Peace geographies and the spatial turn in peace and conflict studies: Integrating parallel conversations through spatial practices. Geography Compass. https://doi.org/10.1111/gec3.12614

Mac Ginty, R., & Richmond, O. P. (2013). The Local Turn in Peace Building: a critical agenda for peace. Third World Quarterly, 34(5), 763–783. https://doi.org/10.1080/01436597.2013.800750c

Mac Ginty, R. (2017). Everyday Social Practices and Boundary-Making in Deeply Divided Societies. Civil Wars, 19(1), 4–25. https://doi.org/10.1080/13698249.2017.1343410

Marquardt, K. L. (2021). Measuring Politically-relevant Identity, With and Without Groups. Varieties of Democracy Institute: Working Paper No. 115. University of Gothenburg.

Miller, Kishi, R., Raleigh, C., & Dowd, C. (2022). An agenda for addressing bias in conflict data. Scientific Data, 9(1), 593–593. https://doi.org/10.1038/s41597-022-01705-8

Mills L., Seets S., Panganiban J. (2021, August 5). AfP's Tech4Peace Series: Introductory Workshop on GIS for Peacebuilding [Webinar]. Alliance for Peacebuilding https://www.youtube.com/watch?v=QDd60lbdRJs

Nemeth, Mauslein, J. A., & Stapley, C. (2014). The Primacy of the Local: Identifying Terrorist Hot Spots Using Geographic Information Systems. The Journal of Politics, 76(2), 304–317. https://doi.org/10.1017/S0022381613001333

Open Geospatial Consortium. (n.d.). Web Map Service. https://www.ogc.org/standards/wms.

Peng, R. (2018, May 24). Context Compatibility in Data Analysis. Simply Statistics. https://simplystatistics.org/posts/2018-05-24-context-compatibility-in-data-analysis/

Pesaresi, Migliara, G., Pavia, D., & De Vito, C. (2020). Emergency department overcrowding: A retrospective spatial analysis and the geocoding of accesses. A pilot study in rome. ISPRS International Journal of Geo-Information, 9(10), 579–. https://doi.org/10.3390/ijgi9100579

Peuquet. (2001). Making Space for Time: Issues in Space-Time Data Representation. GeoInformatica, 5(1), 11–32. https://doi.org/10.1023/A:1011455820644

Pickering, S. (2016). Introducing SpatialGridBuilder: A new system for creating geo-coded datasets. Conflict Management and Peace Science, 33(4), 423–447. https://doi.org/10.1177/0738894215581320

Präger, Kurz, C., Böhm, J., Laxy, M., & Maier, W. (2019). Using data from online geocoding services for the assessment of environmental obesogenic factors: A feasibility study. International Journal of Health Geographics, 18(1), 13–13. https://doi.org/10.1186/s12942-019-0177-9

Raanan M. G., & Shoval, N. (2014). Mental maps compared to actual spatial behavior using GPS data: A new method for investigating segregation in cities. Cities, 36, 28–40. https://doi.org/10.1016/j.cities.2013.09.003.

Raleigh, C., Linke, A., Hegre, H., & Karlsen, J. (2010). "Introducing ACLED: An armed conflict location and event dataset: Special data feature". Journal of Peace Research, 47(5), 651-660. https://doi.org/10.1177/0022343310378914

Raleigh, C., Carew M. B., Carboni A. (2018). Conflict environments and coverage. Clingendael Report. https://www.clingendael.org/sites/default/files/2018-01/Report_Conflict_environments_and_coverage.pdf

Raleigh, C., & Kishi, R. (2019). Comparing conflict data: Similarities and differences across conflict datasets. Armed Conflict Location & Event Data Project (ACLED). https://acleddata.com/acleddatanew/wp-content/uploads/2022/02/ACLED_WorkingPaper_ComparisonAnalysis_2019.pdf

Raleigh, C., & Kniveton, D. (2012). Come rain or shine: An analysis of conflict and climate variability in East Africa. Journal of Peace Research, 49(1), 51–64. https://doi.org/10.1177/0022343311427754

Richmond, O.P. (2009). Becoming Liberal, Unbecoming Liberalism: Liberal-Local Hybridity via the Everyday as a Response to the Paradoxes of Liberal Peacebuilding. Journal of Intervention and Statebuilding, 3(3), 324–344. https://doi.org/10.1080/17502970903086719

Rosvold, E. L., & Buhaug, H. (2021). GDIS, a global dataset of geocoded disaster locations. Scientific Data 2021 8:1, 8(1), 1–7. https://doi.org/10.1038/s41597-021-00846-6

Salehyan, Idean, Cullen S. Hendrix, Jesse Hamner, Christina Case, Christopher Linebarger, Emily Stull, and Jennifer Williams. "Social conflict in Africa: A new database." International Interactions 38, no. 4 (2012): 503-511.

Santillán, P., Edwards, S., Swall, J., & Simmons, T. (2022). Mass grave localization prediction with geographical information systems in Guatemala and future impacts. Journal of Forensic Sciences, 67(1), 112–127. https://doi.org/10.1111/1556-4029.14889

Schutte. (2017). Geographic determinants of indiscriminate violence in civil wars. Conflict Management and Peace Science, 34(4), 380–405. https://doi.org/10.1177/0738894215593690

Singh, S. K. (n.d.). Evaluating two freely available geocoding tools for geographical inconsistencies and geocoding errors. https://doi.org/10.1186/s40965-017-0026-3

Singh, S. K., & Rafiei, D. (2018). Strategies for Geographical Scoping and Improving a Gazetteer. WWW '18: Proceedings of the 2018 World Wide Web Conference, April 2018, 1663–1672. https://doi.org/10.1145/3178876.3186078

Shellman, S. M. (2008). Coding Disaggregated Intrastate Conflict: Machine Processing the Behavior of Substate Actors Over Time and Space. Political Analysis, 16(4), 464–477. http://www.jstor.org/stable/25791950

Smith, H. (2020, March 3). Coordinate Systems: What's the Difference? ArcGIS Blog. https://www.esri.com/arcgis-blog/products/arcgis-pro/mapping/coordinate-systems-difference/

Soifer, H. (2019). Units of Analysis in Subnational Research. In A. Giraudy, E. Moncada, & R. Snyder (Eds.), Inside Countries: Subnational Research in Comparative Politics (pp. 92-112). Cambridge University Press. doi:10.1017/9781108678384.003

Soja, E. W. (1989). Postmodern geographies: The reassertion of space in critical social theory. Verso.

Solinge, D.V., Berens, J., Sambuli, N., Easterday, J., Arthur, P., Panic, B. (2021, September 28). Data for Peace Dialogue: "Do No Harm" In the Digital Age [Webinar]. Center on International Cooperation at NYU. https://www.youtube.com/watch?v=0RHm1knGyjY

Souris, M. (2019). Epidemiology and Geography: Principles, Methods and Tools of Spatial Analysis. John Wiley & Sons, Incorporated. https://doi.org/10.1002/9781119528203.

South Sudan. (2021). Lou Nuer-Dinka Bor-Murle action for peace (Jonglei State). https://www.peaceagreements.org/wview/2378/.

START (National Consortium for the Study of Terrorism and Responses to Terrorism). (2021). Global Terrorism Database (GTD) [Data set]. University of Maryland. https://www.start.umd.edu/gtd

Strauch, R., Keremo˘glu, E., & Weidmann, N. B. (2022). Coding Instructions for the Mass Mobilization in Autocracies Database (Version 4.0) [Codebook]. University of Konstanz, Department of Politics and Public Administration.

Sundberg, R., & Melander, E. (2013). Introducing the UCDP Georeferenced Event Dataset. Journal of Peace Research, 50(4), 523–532. http://www.jstor.org/stable/23441216

Suzuki, A. (2022). Uncertainty in Grid Data: A Theory and Comprehensive Robustness Test. Arxiv. https://arxiv.org/abs/2202.03241

Swift, J., and Wilson, J. (2008). Geocoding Best Practices: Reference Data, Input Data and Feature Matching.

Tollefsen, A. F., Strand, H., & Buhaug, H. (2012). PRIO-GRID: A unified spatial data structure. Journal of Peace Research, 49(2), 363–374. http://www.jstor.org/stable/23141370

UK Data Service. (n.d.). Research Data Management. Retrieved May 24, 2023, from https://ukdataservice.ac.uk/learning-hub/research-data-management/

Uppsala University Department of Peace and Conflict Research. (2021). Policy on Ethical Guidelines for Student Work

Van Baalen. (2018). "Google wants to know your location": The ethical challenges of fieldwork in the digital age. Research Ethics, 14(4), 1–17. https://doi.org/10.1177/1747016117750312

Vogt, M., Bormann, N-C., Rüegger, S., Lars-Erik Cederman, L-E., Hunziker, P., and Girardin, L. (2015). "Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Data Set Family." Journal of Conflict Resolution 59(7): 1327–42.

Weidmann, N. B., and Rød, E. G. (2019) The Internet and Political Protest in Autocracies. Chapter 4. Oxford University Press.

What is Geocoding? (n.d.). Retrieved from https://desktop.arcgis.com/en/arcmap/latest/manage-data/geocoding/what-is-geocoding.htm.

Wise, L., Badanjak, S., Bell, C., & Knaussel, F. (2021). Pandemic Pauses: Understanding Ceasefires in a Time of Covid-19. Political Settlements Research Programme. http://Understanding Ceasefires in a Time of Covid-19

Yusuf, U. (2018, November 13). The difference between "Geocoding" and "Geolocation". Geospatial Solutions Expert. https://umar-yusuf.blogspot.com/2018/11/the-difference-between-geocoding-and.html

Zhukov, Y. M., Davenport, C., and Kostyuk, N. "Introducing xSub: a new portal for cross-national data on Subnational violence." Journal of Peace Research 56, no. 4 (2019): 604-614. https://doi.org/10.1177/0022343319836697

Appendix

Articles that use Geocoded Datasets

The following are other examples of studies within academic literature that use geocoded datasets, listed by region.

Region	Article
Africa	Rustad, Rød, J. K., Larsen, W., & Gleditsch, N. P. (2008). Foliage and fighting: Forest resources and the onset, duration, and location of civil war. Political Geography, 27(7), 761–782. DOI: 10.1016/j.polgeo.2008.09.004
Africa	Raleigh, C., & Kniveton, D. (2012). Come rain or shine: An analysis of conflict and climate variability in East Africa. Journal of Peace Research, 49(1), 51–64. DOI: 10.1177/0022343311427754
Africa	Jensen, C.B., Kuenzi, M.T., & Rissmann, M.P. (2017). Does Crime Pay Enough? Diamond Prices, Lootability and Ethnic War.
Asia	Ali, Rizwan & Khan, Mobushir Riaz & Mehmood, Hannan. (2017). Incidence of Violence Risk Mapping Using GIS: A Case Study of Pakistan. Journal of Geographic Information System. 09. 623-636. DOI: 10.4236/jgis.2017.96039
Asia	Dincecco, Fenske, J., Menon, A., & Mukherjee, S. (2021). Pre-Colonial Warfare and Long-Run Development in India. The Economic Journal (London). DOI: 10.1093/ej/ueab089

Region	Article
Europe	Vogel. (2018). Understanding the impact of geographies and space on the possibilities of peace activism. Cooperation and Conflict, 53(4), 431–448. DOI: 10.1177/0010836717750202
Europe	Darques. (2017). Mapping Versatile Boundaries Understanding the Balkans / by Regis Darques. (1st ed. 2017.). Springer International Publishing. DOI: 10.1007/978-3-319-40925-2
Middle East	Turkmani. (n.d.). Local agreements as a process: the example of local talks in Homs in Syria. Peacebuilding, ahead-of-print(ahead-of-print), 1–16. DOI: 10.1080/21647259.2022.2032941
Middle East	El-Mallakh, Maurel, M., & Speciale, B. (2018). Arab spring protests and women's labor market outcomes: Evidence from the Egyptian revolution. Journal of Comparative Economics, 46(2), 656–682. DOI: 10.1016/j.jce.2017.12.004
Global	Höglund, Melander, E., Sollenberg, M., & Sundberg, R. (n.d.). Armed Conflict and Space: Exploring Urban-Rural Patterns of Violence. In Spatializing Peace and Conflict (pp. 60–76). Palgrave Macmillan UK. DOI: 10.1057/9781137550484_4
Global	Klosek, Bahenský, V., Smetana, M., & Ludvík, J. (2021). Frozen conflicts in world politics: A new dataset. Journal of Peace Research, 58(4), 849–858. DOI: 10.1177/0022343320929726

View Our
Key Findings

info@peacerep.org
PeaceRepResearch
@Peace_Rep_
PeaceRep

www.peacerep.org
www.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, Dialectiq, Edinburgh Law School, International IDEA, LSE Conflict and Civicness 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 the Foreign, Commonwealth and Development Office (FCDO), UK.

PeaceRep: The Peace and Conflict Resolution Evidence Platform
info@peacerep.org | https://peacerep.org | @Peace_Rep_

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