The Anatomy of Resistance Campaigns (ARC) Dataset Codebook: Version 1.0

Charles Butcher, Jessica Maves Braithwaite, Jonathan Pinckney,
Eirin Haugseth, Ingrid Vik Bakken and Marius Swane Wishman

June 15, 2021

1 Introduction

The ARC dataset (version 1.0) contains information on 1,426 organizations that participated in maximalist dissent in African countries from 1990 through 2015 (3,407 country-organization-years in total). ARC includes organizations that participated in violent dissent, such as rebel groups or terrorist organizations, and organizations that participated in non-violent dissent, including political parties, trade unions, and others. The unit of observation is the country-organization-year.

The ARC dataset focuses on *organizations* that participated in acts of *collective dissent* for goals of *maximalist* change. *Organizations* are structures designed to coordinate people and resources – often through collective action – for collective goals (North, 1990; Daft, 1992, 2). The presence of a formal structure (however thin the hierarchy) intended to aggregate individual efforts towards a defined goal separates organizations from broad social categories such as "students," "protesters," or the "working class".

Collective dissent is observable action by multiple people beyond normal institutional procedures for realizing political goals (Tilly, 1978). Collective dissent ranges from demonstrations and strikes to insurgency and terrorist attacks, while excluding actions lacking a clear political goal and everyday or institutional political activities such as lobbying politicians or electoral participation. Organizations engage in collective dissent when they deploy their mobilization infrastructure to encourage individual members to participate in these events.

We define *maximalist* demands as calls for changes in the political structure which, if implemented, would significantly alter the executive's immediate access to state power, the rules with which executives are selected, or the policy or geographic areas for which the executive has the right to make laws. Examples of maximalism include demands that a head of state resign via a non-institutional method, demands for democratization in autocratic settings, demands to enfranchise an excluded social group, and demands for regional or ethnic autonomy or independence.[1] Maximalist demands exclude calls that fall short of altering

[1] A series of borderline demands and their treatment can be found at the ARC project website.

these fundamental aspects of executive power, such as improved human rights protections or changes in public spending. Demands by a disenfranchised group for better human rights protection can be addressed with legislation that typically does not change the process for deciding who holds executive power or who has the right to make laws. Demands for enfranchisement of the same excluded group are maximalist because – if implemented – they would include a new group in the process of deciding who holds executive power.

Only organizations that participated in dissent on the territory of a given country are included in that country's register of active organizations. A Senegalese-based political party that engaged in a protest for democratization in Senegal would be included in the ARC data, while an activist group in London, or South Africa protesting in London or South Africa for democracy in Senegal would not. We discuss the complicated case of the Islamic State (IS) below, but only affiliates of IS that were locally based are counted as actors in a given country. For example, *Jund al-Khilafah* (Soldiers of the Caliphate) is an IS affiliate based in Algeria and for Algeria we would include this organization as an actor in ARC; however, we would not also include *Ansar Bayt al-Maqdis* which was the Egyptian affiliate of IS as an actor in Algeria. This is even though both organizations are – in ARC – considered to be a part of the IS 'front' organization. More details about front organizations and IS can be found below.

2 Collecting the ARC data

Collecting the ARC data involved two steps: (1) identifying organizations that participated in events of maximalist collective dissent and (2) recording information on the features of these organizations. A central value throughout the process was replicability and transparency. At each step we recorded all our coding decisions in RMarkdown files that, when executed, both create the dataset and provide comprehensive documentation of the entire coding process. These Markdown files available on request and we are working to eventually compile them into a publicly available document.

2.1 Identifying Organizations

Dissident organizations were identified by drawing on five events datasets: the UCDP Georeferenced Event Dataset (Sundberg and Melander, 2013), the Social Conflict Analysis Dataset (Salehyan et al., 2012), the Mass Mobilization Dataset (Clark and Regan, 2019), the Armed Conflict Location Event Dataset (Raleigh et al., 2010), the NAVCO 3.0 data covering African countries (Chenoweth et al., 2018). Together these datasets offer a comprehensive catalogue of nonviolent and violent collective dissent across Africa. First, we created a list of *candidate* maximalist events from 1990-2015 by sub-setting on variables in these datasets related to dissident demands and a customized text-matching string.

We examined each candidate event to determine whether participants made maximalist demands, and whether one or more named organizations participated. To accomplish this we located original newswires from which the event was coded, and conducted additional newswire searches in FACTIVA and LexisNexis using a targeted search string. Each of the datasets (except NAVCO 3.0) that we used to identify events of maximalist dissent have unique event ID numbers. Associated event IDs are attached to relevant organization-year observations in the ARC dataset, allowing users to identify the basis by which an organization is included in ARC, as well as to reconstruct aspects of the organization's activity such as its degree of violence or geographic dispersion. Our coding procedures improve upon actor identification in SCAD, MM and ACLED by looking beyond the article that the SCAD/MM/ACLED event was based on by searching for newswires around the event (i.e. for newswires before the event that might feature announcements that certain organizations were planning the event) or after the event for more information as finer details of the event make it into media. Compared to SCAD and MM we also consulted a much broader range of sources including local news sources provided in FACTIVA.

We also included the constituent organizations of "fronts" according to a "three year rule." Fronts are distinct, umbrella organizations coordinating the actions of member organizations. Some datasets like the UCDP treat fronts as unitary actors, but this obscures variation in the preferences and features of member organizations. On the other hand,

treating fronts as decentralized organizational networks for their entire existence can be impractical (and empirically inaccurate). Fronts often become more unified over time or else split apart, but systematically determining the point at which a front ceases to consist of semi-autonomous groups and becomes a single organization is extremely difficult. The National Council of Civil Society Organizations in Guinea (CNOSCG) was formed in 2002 to push for democratization in Guinea by coordinating the activities of NGOs, religious organizations, and trade unions, making it a "front candidate." CNOSCG is included in ARC in 2007 and 2009. However, CNOSCG developed grassroots organizational structures over time with local and regional committees and an executive committee. Based on scarce information it was impossible to determine at what specific point CNOSCG ceased to become a front (if at all) and transitioned to a standalone organization. To resolve this issue, we adopted an arbitrary but empirically informed rule whereby member organizations of a front were added as participants in an event when the tie from that organization to the front was less than three years old. Member organizations were identified in newswire databases, primary and secondary sources, and through an iterative process when information on their features was collected by coders.

For example, the "Allied Front for Change" (FAC) was formed in Cameroon by four political parties (the SDF, MDP, MP and UDF) in October 1994 and organized protests in that same month. When news sources identified the FAC as a participant in an event, the FAC and the four parties were included in ARC as participants because the political parties had been members of the front for less than three years. If, hypothetically, the FAC formed in 1990 and participated in a protest in 1994, ARC would include the FAC as a participant but not the four political parties (plus we include any other participants identified in the newswires). ARC never includes *fewer* organizations than those explicitly identified in the news sources around an event, but the front rule means that we sometimes included *more* organizations when the tie between the front and a constituent group was less than three years old. The only exceptions are when organizations clearly merged and ceased to exist independently, or when an organization left a front. For example, the Trade Unions Congress

of Swaziland (TUCOSWA) is a merger of the Swaziland Federation of Trade Unions (SFTU) and the Swaziland Federation of Labour (SFL). The SFTU and SFL deregistered as trade unions after the merger and ceased to exist.

In most cases, the age of ties between constituent organizations and the front are the same as the age of the front itself, and most constituent organizations are subsumed into the front after the front endures for more than three years. However, organizations can join pre-existing fronts, and for the first three years of their front membership these joiners are added as distinct participants to events involving the front. For example, the National Constituent Assembly (NCA) in Zimbabwe was formed in late 1997 by religious groups, the Zimbabwe Congress of Trade Unions (ZCTU) and other civil society organizations. We do not include the ZCTU as a separate participant in NCA events after 2000 (unless it is explicitly identified as having participated independent of NCA), but we do add Women of Zimbabwe Arise (WoZA) as a participant when the NCA participates from 2004-2007 because WoZA joined the NCA in 2004 and thus the tie is less than 3 years old during this period.

While any such threshold is arbitrary, we argue this strikes a balance between theoretical and practical concerns. The UCDP amalgamation rule might be practical for some rebel groups, but certainly obscures too much when examining the broad fronts that often make up nonviolent movements. However, to disaggregate all fronts would be impractical and would overlook fronts' typical move towards unity over time. In an ideal world we could precisely identify when fronts ultimately merge into a unitary organization, but in our experience information is simply too sparse to achieve this. As such, we rely on an arbitrary cut-off, but one that we believe balances theoretical and resource demands.

It would be possible to impute forwards the participation of organizations in a front beyond the three years that they were members. For example, users could assume include members of the "Allied Front for Change" (discussed earlier), formed in 1990, as participants from 1994 onward by including the same member organizations from 1993. However, there are two disadvantages to this: (a) we don't know if these organizations were still members

and, (b) organization-level features would also have to be imputed forwards if users wanted to retain information on organization-features. Both of these disadvantages could introduce error.

The rules for coding front participants means that some organizations that were members of relatively new fronts but did not participate in protests, or played only a peripheral role, may be included in ARC. However, we argue that this risk is outweighed by the benefit of including organizations that are often elided by the news media, such as local human rights organizations, women's organizations and youth groups, that are often crucial to individual mobilization. Since front participants are identified through newswires *and* primary and secondary sources, the front rule makes our inclusion criteria less dependent on what journalists and the news media typically cover.

2.2 Coding Organization Features

This identification process produced a list of organizations linked to events in these five datasets. From this we generated a list of dissident organization-years, and distributed that list among a team of coders who recorded information on organizational features in a given year. While some variables remain constant for all observations of an organization (such as its date of origin), other variables are dynamic. Table 7 summarizes our main organizational feature variables.[2] More information on each variable can be found below.

ARC also includes information on two types of ties between organizations: fronts and alliances. Front ties exist between a constituent organization and a "higher-level" organization (a front) when the constituent organization is formally a member of the front, or its leaders participate in the front's leadership.[3] Organizations identified by the 'three year rule' discussed earlier have front ties to the main front. When a front and a member organization participate in a given year independent of the front rule, they still have a front tie. The

[2]The full codesheet can be found on the ARC project website with specific instructions for coding each variable.

[3]In some cases, fronts themselves become constituent organizations in additional higher-level fronts. In this case, we only include ties from constituent organizations to the closest-level front in the hierarchy.

Table 1: Organization-level variables

Variable	Description	Format
Type	Categorization of organization type	Categorical
Birthdate	Date organization was founded	Date: 'dd-mm-yyyy'
Origins	How organization formed	Categorical: (Splinter, Merger, Other)
Goals	Primary organization goals	Open text
Size	Membership size in year	Numeric
Size Estimate	Approximate size	Ordinal
Leadership	Leader name/gender	Open text
Leadership Tenure	Date leader assumed position	Date: 'dd-mm-yyyy'
Leadership Ties	Did leader serve at a high level in previous governments?	Categorical: (Yes/No)
Social Base	Main social group(s) in organization	Open text
Social Media	Extent of social media use	Categorical: (None, Some, Significant)
State Rel.	Relationship with state at t-1	Categorical
Formal Ties	Ties with other active organizations	String: Organization IDs
Structure I	Clear leadership/decision making structure?	Categorical: (Yes/No)
Structure II	Characterised as 'decentralised'.. etc.	Categorical: (Yes/No)

ZCTU, for example, is a participant in protests against the Mugabe regime in 2002, 2003, 2005 and 2007. The NCA also participates in these years. The ZCTU is a member of the NCA through the 2000s and the ZCTU and the NCA have a front tie for these years in ARC *even though the ZCTU's participation was not identified on the basis of the front rule.*

Alliance ties are included when two or more organizations declared the coordination of their resistance activities, or sources indicated that organizations coordinated efforts, but they did not form a standalone organization (a front) to manage this coordination. Fronts and their constituent organizations can have alliance ties with organizations outside the front. For example, in Malawi in 1993, the Public Affairs Committee (PAC, a front of CSOs and religious groups) allied with the Alliance for Democracy (AFD, a new political party), which was not formally a part of the PAC. Users can assemble "alliance-pairs" and organization-dyads with these front and alliance variables to explore factors driving inter-organizational ties.

Figure 1 illustrates these different types of ties through an example network. The organization at the bottom-center of the network has alliance ties to two other organizations, and is also a member of a front. That front in turn is a member of a second front.

Ties in the ARC data are undirected edges as we have not explicitly conceptualised or theorised directionality in terms of front ties. For Figure 1 arrows moving from constituent organizations to fronts best captured the fact that constituent organizations form fronts (and generally not the other way around), but directionality should not be inferred from the data.

Figure 2 shows the types of networks that this process records, tracking networks of organizational participation in maximalist dissent in Egypt in 2011.

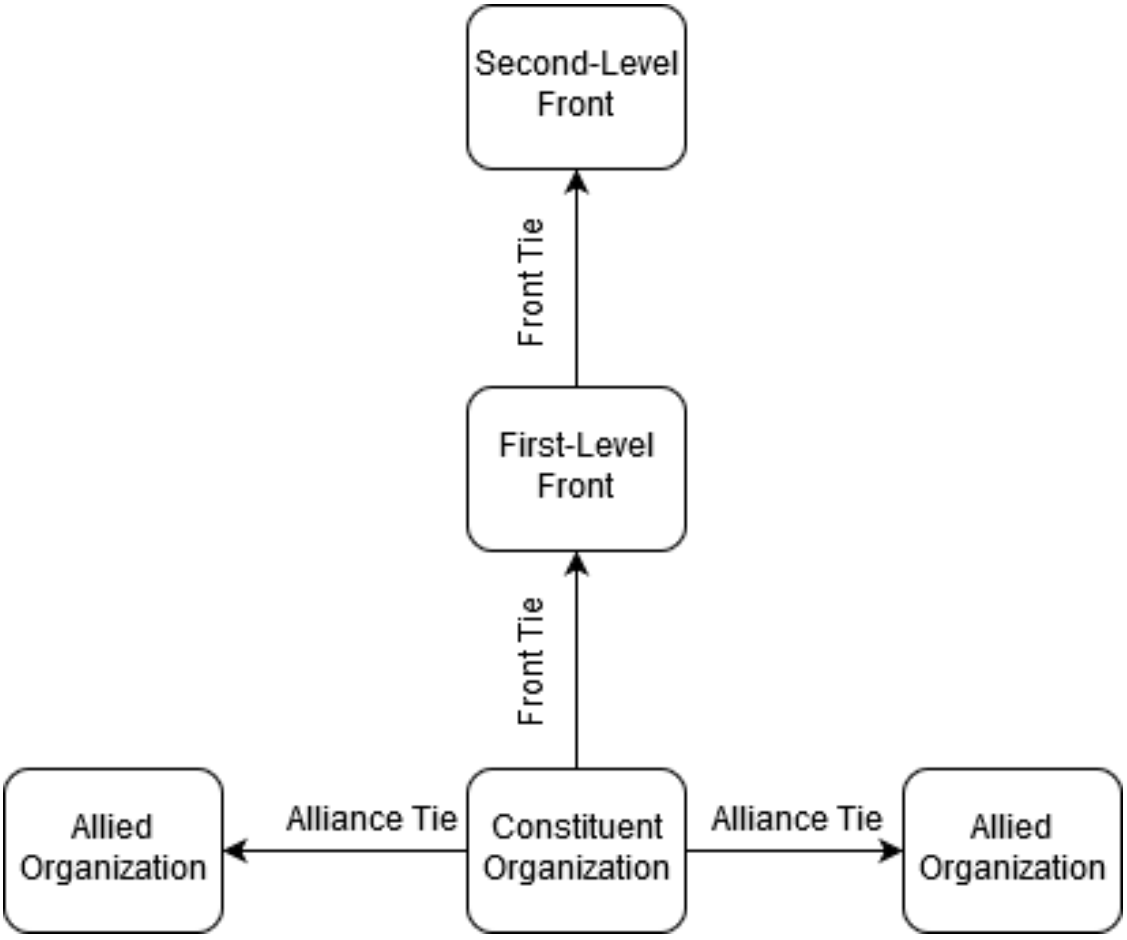

Figure 1: ARC Ties Example

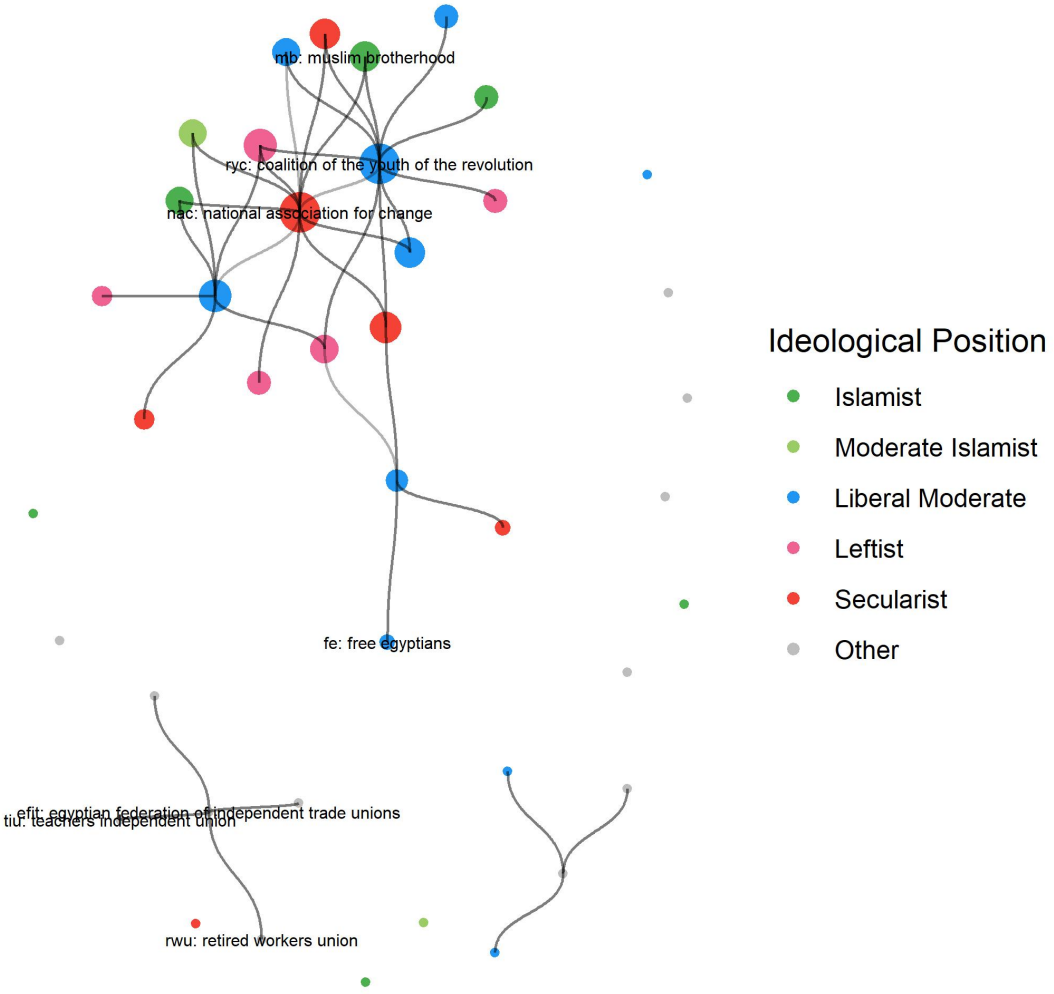

Figure 2: Egypt 2011

There are a number of approaches to expanding the ARC data into an organization-dyad format to study network formation or the correlates of network ties. To create a "full" organization-dyad dataset, one could take each unique organization observed in ARC (or within in certain country or timeframe) and expand it to a data frame with an observation for each unique-pair of organizations in the data (or in a given year, or country-year). Researchers could then create a unique organization-pair ID (or organization year pair ID) by concatenating the existing organization IDs. For pairs that form alliances (i.e. an "alliance-

tie dataset"), ARC can be transformed into an alliance-dyad year format by splitting the "alliances" variable (separating on the comma) and then expanding the data to create a unique observation for each alliance pair. This "alliance" data could then be merged with the full organization-dyad data frame by organization -pair IDs.

3 General Coding Rules

The section below contains detailed information on the coding of each individual variable. However, a few general rules that we applied across all variables are important to clarify first:

Missing Data: All variables with missing data are coded NA. This applies to both numeric variables and character-based variables.

Temporal Inference: Since our unit of observation is the organization-year, for most variables we only include information that we could directly link to the organization in the year in question. For instance, we only code the org.size variable if we have a specific estimate of the organizations size in the year being coded (not, for instance, if we have an estimate from other years before or after). Users of the data can use the existing information to make reasonable inferences based on their preferences.

Sourcing and Inter-coder reliability: We instructed coders to follow a specific sequence of checking sources when coding variables. First, for rebel groups coders consulted the Uppsala Conflict Data Programs Conflict Encyclopedia. For nonviolent groups, coders consulted the Political Handbook of the World, the Historical Dictionary for the country in question, and the Global Nonviolent Action Database. For any variables they were unable to find information on in these initial sources, they then consulted the reports available on the Refworld website. For any remaining variables they consulted newswires through the FACTIVA media aggregator. For any remaining variables after this they consulted scholarly articles or other secondary sources.

Conflicting Sources and Evidence: We instructed coders to keep detailed notes on the evidence they used to code variables, with specific references and, if necessary, an

explanation of their logic in choosing between conflicting sources. These detailed notes are kept in individual "codesheets" for each observation. We hope at a later date to be able to make all of these codesheets publicly available on an ARC website. For now, if you have questions about any specific coding and would like access to an individual codesheet, dont hesitate to ask us.

4 Specific Variable Descriptions

4.1 Country Name (country.name)

The name of the country where the organization engaged in maximalist contention.

4.2 Year (year)

The year the organization engaged in maximalist contention.

4.3 Organization Name (org.name)

The name of the organization, generally translated to English.

4.4 Other Names and Acronyms (org.name.other)

Other names and acronyms sometimes used in sources to refer to the organization.

4.5 Organization ID number (org.id)

The organization's unique Anatomy of Resistance Campaigns dataset ID number. These take the form of a six-character identifier with a dash in the middle. The first three characters identify the organization's home country. The last three characters are a unique numeric code for that organization. For example, Boko Haram's organization id is NIG-001. The African National Congress in South Africa's ID is SAF-005.

4.6 Organization Type (org.type)

Specifies the type of organization, based on an eight-level categorization. Organizations that fall into more than one category will have multiple categories listed, separated by a comma. The categories are as follows:

Table 2: Organization types

Type	Description
1.Political Party	Political parties are groups whose primary objective is to achieve political power and who are formally organized to achieve this goal. They achieve this goal primarily through attempting to win elections. However, political parties can exist in countries that either do not have elections, and in countries where political parties are illegal.
2.Trade Union	Trade unions are organizations made up primarily of workers whose primary mission is to promote workers' interests. Trade unions are often limited to a single industry (such as a Teachers' Union or a Steelworkers' Union) but need not be so. Trade unions need not be formally recognized by the government or by international trade union bodies in order to be coded as a trade union
3. Formal Religious Community	These are organizations whose primary mission is to organize religious practice. Note that this does not include groups who may have a religiously-based identity but whose primary mission is not about religious practice (i.e political parties with religious goals).
4. Student/Youth group	Organizations whose membership is exclusively or primarily students or youth and who are organized around that identity
5. Front	Fronts are organizations made up of other organizations. Often these are short-term expedients, for instance when various political forces create a front to coordinate a shared protest campaign. Many groups that may have "Front" in their name are not necessarily fronts, and many fronts do not include the word "front" in their name.
6. CSO not otherwise mentioned	Civil society organizations are any group in which private individuals come together to pursue a common goal outside of state structures and not primarily in pursuit of financial gain. Most nonviolent groups that don't fall into categories 1-5 fall into this category.
7. Rebel Group	A rebel group is an organization that seeks to achieve its goals through the application of violent force. Rebel groups are armed, have at least some form of organizational structure, and endure over time
8. Other	Any organization that does not reasonably fall into one of the above categories should be coded here. This category should be used sparingly, as almost all groups should fall into one of the first seven categories.

4.7 Organization Birth Date (org.birth)

The founding date for the organization, coded with as much specificity as possible. The variable is coded following the format dd-mm-yyyy. If only the year or the month and year are known, only these portions of the date are filled in. For instance, if the coder discovers that an organization was founded in June 1986, but did not know the exact day, they would code the variable as dd-06-1986.

Some groups may be dissolved and then re-form before the observation year. In this case, the start year was coded as the year the group was re-formed. For example, if the observation year was 1995 and there was information that the group was originally founded in 1960, but was then dissolved in 1965 and also came back into existence in 1975, 1975 would be coded as the start year, not 1960.

4.8 Organization Origins (org.origins)

Three level categorical variable capturing whether the organization was a splinter from a previous organization, a merger of two or more pre-existing organizations, or a newly-formed organization. The levels of the variable are as follows.

Table 3: Organization origins

Code	Start type
1	Splinter: A splinter is an organization made up of "members who broke away from a previously existing group" (Lounsbery 2016). For instance, the CNDD-FDD rebel group in Burundi was a splinter group because it was created when the military wing of a pre-existing rebel group (the CNDD) broke away to form their own group.
2	Merger: A merger is any organization that is formed from multiple pre-existing organizations coming together to form a new group. Fronts are always mergers, but other groups may also be mergers. For instance, a political party that was created when two smaller political parties joined forces would be a merger.
3	Other: the organization was not a splinter or a merger.

Some groups have elements of both splintering and merging in their origin stories. For example, the CNT in Niger is a union that was formed by several small groups breaking away from a preexisting union and then coming together to form a new union. In these cases, if the splinters had names and existed independently for a meaningful period of time, the organization was coded as a merger. If, however, the new organization quickly came into being after breaking away from the old organization, it was coded as a splinter.

4.9 Organization Goals (org.goals)

Open text field listing short summaries of the publicly-stated goals and ideology of the organization, separated by commas. For instance, a movement seeking a proletarian revolution and independence from a central states entry could be: Marxist,separatist.

4.10 Organization Size (org.size, org.size.low, org.size.high)

A precise numeric estimate of the number of formal members of a group. These must be estimates of a specific number of members cited in a reliable secondary source that refer to the observation year. Organization size is not equivalent to participants in a protest, number of likes or followers on Facebook, or other indicative indicators of organization size.

When more than one estimate is discovered, the coder should put the highest estimate found in the org.size.high variable and lowest estimate found in the org.size.low variable. Coders should put the most reliable estimate (based on their examination of the case) in the org.size variable. If they are uncertain which estimate is more reliable, they should put the average of the org.size.low and org.size.high variables in the org.size variable.

Note: most observations of org.size, org.size.high and org.size.low are missing.

4.11 Organization Size Estimate (org.size.est)

The organization size estimate is a categorical variable containing a rough order-of-magnitude estimate of the number of members of the organization. Coders may use any data that helps inform reasonable inferences about the number of members in the organization to help gener-

ate this estimate (such as participants in the protest, social media followers, election results). Coders must describe their logic in coming to this inference on the codesheet.

Table 4: Organization Size Estimate

Category	Numerical Estimate
1. Very small	0 – 100 members
2. Small	101 – 1,000 members
3. Medium	1,001 – 10,000 members
4. Large	10,001 – 100,000 members
5. Very large	100,001 – 1,000,000 members
6. Over a million	> 1,000,000 members

4.12 Organization Leader (org.lead)

The name of the most important decision maker(s) for the group during the year being coded. If possible to determine, this is the formal leader of the group based on established group rules and institutions. If this is impossible to determine, or such established rules and institutions do not exist, then the individual that the preponderance of sources refer to as the group's leader or primary power-holder. If sources are unclear on which of a small number of individuals is the primary leader, multiple leaders may be coded, with names separated by commas.

4.13 Organization Leader Gender (org.lead.gender)

The publicly-identified gender of the leader, with M signifying male and F signifying female. If multiple leaders were identified in the org.lead variable, then the publicly-identified gender of each leader separate by commas (for instance, M,F,F).

4.14 Leadership Tenure (org.lead.start)

The date on which the leader during the year being coded became the group leader. As with the organization birth date, the variable is coded following the dd-mm-yyyy convention, with as much of the date filled in as possible.

4.15 Leadership Ties (org.lead.ties)

Binary variable capturing whether the group leader at some point in the past held a formal role at a high level of government. We define this as an individual in an official government position who has a regular and meaningful impact on national-level policy. This includes members of parliament, for instance, but not local government officials. For former members of the military, this is all individuals with a rank of Colonel (or equivalent) and higher. The variable is coded 1 if the leader had such a role in the past, and 0 if they did not.

To be coded as a 1, the leader themselves must have had a direct, formal role in government. They cannot have simply been connected to government through family or professional connections. So, for instance, if the leader's uncle was a minister, but the leader never had a formal role in government, code the variable as a 0.

4.16 Social Base (sb)

Open text field capturing succinct descriptions of the primary social groups from which the organization mobilized, with individual descriptions separated by commas (e.g. students,workers,intellectuals).

Note that a social base is a broad category of society, not a specific organization. So, for instance, if the group is a splinter, the parent group is not one of the social bases.

4.17 Mobilization and Social Media (social.media)

Three-level categorical variable that measures the organization's use of social media for mobilization. The variable is only coded beginning in 2006, since this is the year in which

the major social media platforms (Facebook and Twitter) opened up to the general public. For all years prior to 2006 the variable is coded as missing. The levels of the variable are as follows:

Table 5: Social media

Category	Description
0	No social media for discovered for this organization in this year.
1	Some social media exists, but no clear evidence of large mobilization.
2	Clear evidence of large mobilization through social media.

4.18 Government Relationship (state.rel)

Categorical variable indicating the organization's relationship to the government in the year prior to the year being coded (so, for instance, in a codesheet for 1991, this variable refers to the relationship in 1990). If there was a significant shift in a group's status during the course of a year (for instance, it goes from being legal to being banned), then this variable was coded with the group's status as of December 31st (in the previous year). The levels of the variable are as follows. Formally included organizations are those with formal, public alliances to the regime in the year before the observation year or organizations that were formally incorporated into the regime. Examples include party-affiliated trade unions, as well as political parties with cabinet or ministry positions (regardless of how undemocratic the regime was in practice). Informally included organizations or privileged organizations are those with regular, explicit access to state power or officials that are not afforded to other institutions of the same type. Neutral organizations are neither included in the regime bargain nor banned. They have a de jure legal existence that is respected in practice. Banned organizations are those that are de-jure illegal. Organizations that arise during the year are coded as not being in existence.

Table 6: Relationship to the regime

Category	Description
1	Formal inclusion in the regime.
2	Informal inclusion in the regime
3	Legal
4	Banned/Illegal
5	Not in existence
6	No significant interactions with the state

4.19 Alliances (alliances)

A list of the organizations with which the organization in question had a formal relationship, with individual organizations separated by commas. We include two kinds of relationships: alliances and fronts. A coder includes an alliance relationship when two organizations make public statements that they will coordinate their activities towards achieving a shared goal. The coder includes a front relationship when an organization participates in creating a new organization to coordinate the activities of multiple organizations (a front), or when the organization being coded is itself a front. If the organization being coded is a front, then we code relationships to each one of the constituent members of the front. If the organization being coded is a member of a front we code that organizations relationship to the front, but not a relationship to the other organizations within the front.

For example, if two political parties: The Democracy and Development Party and the Alliance for Change announce that they are coordinating a resistance campaign against the government, then we code an alliance relationship between them. If, however, they announce that they have created a new organization: The Front for Political Transformation, then we would code each as having a relationship to the front, but not a relationship directly to each other. The front would be coded as having relationships to both.

We include the organization ID numbers in this field. Only alliances to other active organizations in ARC are coded.

4.20 Alliance Types (alliances.type)

This variable lists which of the two relationships (front or alliance) each of the organizations identified in the previous variable had with the organization being coded, with each relationship separated by a comma. Front relationships are indicated by a 1 and alliance relationships are indicated by a 2. If there are no relationships observed, then the variable is coded 0.

The order of relationships should correspond to the order in which the specific organization ID numbers were entered into the alliance variable. For instance, if the organization being coded had a front relationship with FOO-001 and an alliance relationship with FOO-002, and the alliances variable had been coded in that order (i.e. FOO-001,FOO-002), then the alliances.type variable should be coded 1,2.

4.20.1 Example coding from coding instructions

You are coding a political party in Chad called the Party for Amazing Democracy in Chad in 1995 (PADC, org-ID CHA-666). You come across a news article that says that that the leader of the PADC met with the leader of another political party, the Front for Achieving Perfect Democracy Everywhere (FAPDE, org-ID: CHA-999), and announced that their parties will be joining together to organize a general strike demanding that the president step down from power. This would be an alliance relationship, so you would put the org-id for the FAPDE - CHA-999 - in the alliances variable, and a 2 in the alliances.type variable.

Later that year the PADC decides that their alliance with the FAPDE isn't enough, and they decide to create a new organization called the Ultimate Front of All Opposition (UFAO, org-ID: CHA-888). Joining them in UFAO are a trade union: the Widget Workers of Chad (WWC), and a student group: Students for an Incredible Chad (SIC). You would now add UFAO's org-ID to the alliances variable, and would add a front relationship to the alliances type variable. You would not add the org-IDs for WWC or SIC because these organizations only have a relationship with PADC through their shared membership in the front.

The final coding of these variables would look like this:

cha.padc.1995$alliances <- 'CHA-999,CHA-888'

cha.padc.1995$alliances.type <- '2,1'

4.21 Decision-making body (org.struc.a)

Did this organization have a formal decision-making body/individual (for example, an executive committee, president)?

You should code this variable as a 1 if there is any kind of formal hierarchy of decision-making evident in the organization. For instance, if a group claims to be "leaderless," but elects a standing committee that makes authoritative decisions, then you would code the variable as a 1.

4.22 Decentralised structure (org.struc.b)

Binary variable capturing whether the organization in question was centralized or decentralized. Coders code this as a 1 if sources refer to the organization as a confederation, federation, loose alliance, umbrella organization, or decentralized and code it as a 0 otherwise.

4.23 Number of associated events (mmd.n, acled.n, ged.n, scad.n, navco.n)

Numeric variable counting the number of events that this organization was associated with in the country-year for each of the datasets that we used to identify organizations from. Specifically these are mmd (Mass Mobilization Dataset, (Clark and Regan, 2019)), acled (Armed Conflict and Location Dataset, (Raleigh et al., 2010)), ged (Georeferenced Events Dataset, (Sundberg and Melander, 2013), scad (Social Conflict Analysis Dataset, (Salehyan et al., 2012), and navco (NAVCO 3.0 dataset, (Chenoweth et al., 2018)).

4.24 ID numbers of associated events (mmd_id, acled_id, ged_id, scad_id, navco_id)

Specific event identifiers from each of the datasets that we used to identify organizations from. Identifiers are unique to each event datset and are separated by a comma.

5 Error Reduction Strategies

In addition to the common codesheet with specific instructions on how to code the variables, we maintained a website through the project with more detailed coding instructions for the identification and coding organization-features stage that included more detailed information on each variable and more specific coding instructions. This Wiki site contained information on sources to be followed, how to code constituents of fronts, inferring information from the past and future, what to do with missing data and conflicting sources and an FAQ. It can be accessed here `https://sites.google.com/view/anatomy-resistance-campaigns/home`.

We also held weekly coding meetings with the research team where we could ask questions and respond, which we believe helped generate a shared understanding of the main concepts and how they should be coded.

We conducted one round of inter-coder reliability tests which we intended to use for training purposes and for identifying areas where coders had shared understanding of concepts and operationalizations and where we needed to develop additional rules or do additional training. The round of tests were very useful for identifying areas where the original coding rules were unclear or ambiguous. Eleven different coders completed five code-sheets each (5 country organization-years) that were randomly selected. The exercise was used to identify areas of agreement and disagreement across coders and provide follow up training as a group. Most variables had generally good agreement (birth-date, type, origins, goals, social base and leadership). The tests revealed problems with interpretation of the leadership ties variables and the state-relationship variables which we clarified and followed up by adding more specific coding instructions on both of these variables, which can be found here and

on the website. The most serious problems were with the size estimate variable, which coders disagreed significantly on. However, coders were generally in agreement on whether the organization was bigger than 10,000 members (i.e. bigger or smaller than a 3 in the codesheet). We would recommend that users consider breaking this down into a simpler categorical indicator potentially small (1-9999 members), large (10,000 99,999 members) and very large (100,000 +) members. We also addressed some of these problems by clarifying the type of information that could be used to infer size estimates. This information is in the codebook. All coders were instructed to revise coding that they had already done in light of the discussions.

In addition to the above, the alliances and fronts variables presented the most problems for coders. These variables have be been systematically reviewed by the authors to ensure consistency for this variable across cases.

6 Specific cases: Islamic State

The Islamic State (IS) illustrates our 3-year coding rule for fronts, and is an unusual case as an example where the front organization is foreign based. First, we treat IS as a 'front' and most of the ties from local insurgent groups to IS as 'front' ties because these groups pledge to be a 'part of' or to act on behalf of IS. This is in contrast to ties to Al-Qaeda which are most often pledges of 'support' to Al-Qaeda not pledges to join Al-Qaeda. We treat these ties to Al-Qaeda as 'alliance' ties. Some groups such as Ansar al Sharia in Tunisia pledge only support for IS and Boko Haram only 'joins' is in 2015, after pledging support in 2014. In these cases we code an 'alliance' tie to IS.

The only case where we treat IS as an independent actor in Africa over the 1990-2015 period is in Libya in 2015, where there was evidence that IS fighters were direct IS members from abroad, not local organizations that pledged to join IS.

Ties to IS start at different points in time for different insurgent groups, sometimes within the same country. We treat these front-ties with the same 3-year rule as other front ties. If the local insurgent group declares that it is part of IS then we treat the local insurgent

group as an independent actor with a front tie to IS for 3-years, after which it becomes a part of IS and IS becomes the actor. All the specific cases of ties to IS are detailed below in Table 7.

Table 7: Ties between insurgent groups and Islamic State

Organization Name	Org ID	Active Countries	Relationship to IS
soc: soldiers of the caliphate	ALG-025	Algeria, Tunisia	Front tie: 2014-2015
boko haram: jama'atu ahlis sunna lidda'awati wal-jihad	NIG-001	Nigeria, Chad, Cameroon, Niger	Date: Alliance tie: 2014, Front tie: 2015
abm: ansar bayt al-maqdis	EGY-002	Egypt	Date: Front tie: 2014-2015
mssi: islamic youth shura council	LIB-046	Libya	Date: Front tie: 2015
as: ansar-al sharia	TUN-031	Libya	Date: Alliance tie: 2013-2015
is: islamic state	IRQ-001	Libya	Front

7 Extracting goals

Users can create variables that capture an organization's goals or it's social base with text matching in `STATA`, `R` or `Python`. For example, the following word stems can identify the vast majority of organizations with separatist goals in the ARC dataset when used on the `org.goals` variable: `secess,autonom,separatis,decentral,federal,selfdetermin, independen,irreden,regional,territor`.

References

Chenoweth, E., Pinckney, J., and Lewis, O. (2018). Days of rage: Introducing the NAVCO 3.0 dataset. *Journal of Peace Research*, page Online First.

Clark, D. and Regan, P. (2019). Mass Mobilization Protest Data. Harvard Dataverse: https://doi.org/10.7910/DVN/HTTWYL.

Daft, R. L. (1992). Organization theory and design, St. *Paul, MN: West Publishing Company*.

North, D. C. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge University Press, Cambridge ; New York, NY.

Raleigh, C., Linke, A., Hegre, H., and Karlsen, J. (2010). Introducing ACLED: An Armed Conflict Location and Event Dataset: Special Data Feature. *Journal of Peace Research*, 47(5):651–660.

Salehyan, I., Hendrix, C. S., Hamner, J., Case, C., Linebarger, C., Stull, E., and Williams, J. (2012). Social Conflict in Africa: A New Database. *International Interactions*, 38(4):503–511.

Sundberg, R. and Melander, E. (2013). Introducing the UCDP Georeferenced Event Dataset. *Journal of Peace Research*, 50(4):523–532.

Tilly, C. (1978). *From Mobilization to Revolution*. Routledge, Boston, MA.