Rebels' Armament Dataset (RAD) Codebook

Version 1.0

Andreas Mehltretter, Paul W. Thurner, Oliver Pamp, Paul Binder

September 30, 2024

1 Introduction

The Rebels' Armament Dataset (RAD) provides detailed information on the armament of non-state actors/rebel groups fighting in state-based conflicts. The data was collected systematically by a multi-stage research process using a variety of sources, such as the NISAT Document Library, the UCDP Conflict Encyclopedia, and a broad range of conflict literature. In addition, an extensive media analysis using Google as well as the Nexis research database was conducted. The retrieved evidence was coded in a standardized form and each type of weapon was assigned to pre-defined categories, which are described in chapter 7. More details on the process of data collection and processing as well as limitations and caveats can be found in (Mehltretter et al. (2023)).

The three following features of this dataset deserve special emphasis in order to fully grasp its structure and content.

1. Two different versions of the Rebels' Armament data are available. The main dataset only includes group-level variables (referred to as "Group-dataset" in the remainder of this document). Hence, the unit of observation is the rebel group. We started by searching for information about armament for 345 rebel groups. The variable 'Group_Information_available' indicates the 269 groups for which we were able to find at least one specific piece of evidence regarding their armament. The second dataset includes time-series, group-year level variables (in the following "Groupyear-dataset"). It encompasses 1,343 observations between 1989 and 2020. We only cover rebel groups during active conflict-years as defined by the UCDP/PRIO Armed

Conflict Dataset version 22.1 (Davies et al., 2022). Please note the uncertainties in attributing evidence to specific time spans, as described in (Mehltretter et al. (2023)) and section 6 of this codebook.

2. The data on rebels' armament is presented at three different levels of aggregation. At the highest level of aggregation, the total armament (concerning all types of weapons) of groups is measured. At the medium level of aggregation, the measurement of rebel groups' arms stocks is dis-aggregated into the five categories: *major conventional weapons*, *light weapons*, *small arms*, *explosives*, and *others*. The definitions of these categories are given in section 7. Finally, at the lowest level of aggregation, information for 14 specific distinct subcategories of weaponry (such as *landmines*, *missile, rocket and grenade launchers*, *rifles and shotguns*, and *tanks*) is provided. Note that not all levels of aggregation are included in both datasets.[1]

3. We developed ordinal as well as metric measures of rebel groups' armament. In tables 2 and 4, the scale of each variable is indicated in the third column. The metric variables presented in this codebook are lower bound estimates, only including entries of evidence with precise or imprecise[2] quantitative information on the amount of weapons recorded. Hence, data entries containing no specific information on amounts are excluded from the calculation of these variables. In order to fully utilize all the information we have compiled, we assigned the median amount of the respective weapons category to these entries before adding them to the total armament measures in the respective categories, and transformed the results into ordinal variables. Hence, these measures are necessarily afflicted with additional uncertainty but give a more complete picture of a group's armament in comparison to the lower bound estimates. These ordinal variables are the cornerstone of the Rebels' Armament Dataset and account for the vast majority of the variables presented in the codebook. All ordinal variables employ the same 5-point scale, ranging between 0 and 4 with 0 indicating no evidence of possession of the given weapons category and higher values denoting larger amounts.

To ensure clarity, the type of variable is indicated by their respective prefixes: 'Group_'

[1]Details on the exact coverage of the datasets with regard to the different levels of aggregation are given in sections 3 and 4.

[2]Imprecise amounts include entries with amounts such as "several", "numerous" or "dozens".

indicates group-level variables in the Group-dataset while 'Groupyear_' refers to the group-year level variables in the 'Groupyear-dataset. The prefixes are chosen in order to allow for merging the two datasets without creating any confusion.

The remainder of this codebook describes the specific variables presented in the RAD. It is structured as follows. Section 2 presents identifier variables while chapters 3 and 4 present the RAD variables, their descriptions, and scales for the Group-dataset and Groupyear-dataset. Finally, chapter 7 defines and describes the weapons (sub-)categories at the different levels of aggregation in more detail.

2 Identifier Variables and General Information

Table 1 depicts the variables that identify each observation, which are based on UCDP variables (cf. Davies et al., 2022), as well as variables that present general information on the rebel groups.

Table 1: List of Identifier Variables and Variables Capturing General Information

Variable name	Content
Group_UCDP_ID	UCDP ID of nonstate actors (cf. Pettersson, 2022, 7)
Year	Year of observation (only included in the Groupyear-dataset)
Group_UCDP_Name_short	Short name or abbreviation of nonstate actor (cf. Pettersson, 2022, 7)
Group_UCDP_Name_long	Full name of nonstate actor (cf. Pettersson, 2022, 7)
Group_UCDP_Earliest_year	Earliest observed year of the group relevant for the data collection according to the UCDP data (cf. Davies et al., 2022; Gleditsch et al., 2002)
Group_UCDP_Latest_year	Last observed year of the group relevant for the data collection according to the UCDP data (cf. Davies et al., 2022; Gleditsch et al., 2002)
Group_Information_available	Dummy, 1 if there is any information on the group available, 0 else (only included in the Group-dataset, groups without any information excluded from Groupyear-dataset)

3 Variables in the Group-dataset

Tables 2 and 3 describe the measurements and variables of rebels' armament at the group-level. As described above, the metric lower bound variables are the sums of the amounts of weapons in a given category. The ordinal lower bound measures are based on both these metric variables as well as additional entries with imprecisely given amounts. See section **??** for details on the aggregation of multiple entries and values into the overall value for a group in a specific arms category.

The number of entries that are used for the calculation of these variables is given by the variable 'Group_Evidenceentries_qnty'. Table 2 covers the medium level of aggregation and some more general variables, such as numbers of transfers and potential countries of origin, while table 3 displays variables at the lowest level of aggregation.

Table 2: List of Variables Included in the Group-dataset (at the Medium Level of Aggregation)

Variable name	Content	Scale
Group_Smallarms	*Small arms* armament per group	Ordinal (0-4)
Group_Lightweapons	*Light weapons* armament per group	Ordinal (0-4)
Group_MCW	*Major conventional weapons* armament per group	Ordinal (0-4)
Group_Explosives	*Explosives* armament per group	Ordinal (0-4)
Group_Other	*Other* weapons per group	Ordinal (0-4)
Group_SA_lowerbound_metric	Lower bound estimate of *Small arms* armament per group	Metric
Group_LW_lowerbound_metric	Lower bound estimate of *Light weapons* armament per group	Metric
Group_MCW_lowerbound_metric	Lower bound estimate of *Major conventional weapons* armament per group	Metric

Group_Expl_lowerbound_metric	Lower bound estimate of *Explosives* armament per group	Metric
Group_Other_lowerbound_metric	Lower bound estimate of *Other* weapons per group	Metric
Group_SA_lowerbound	lower bound estimate of *Small arms* armament per group	Ordinal (0-4)
Group_LW_lowerbound	lower bound estimate of *Light weapons* armament per group	Ordinal (0-4)
Group_MCW_lowerbound	lower bound estimate of *Major conventional weapons* armament per group	Ordinal (0-4)
Group_Expl_lowerbound	lower bound estimate of *Explosives* armament per group	Ordinal (0-4)
Group_Other_lowerbound	lower bound estimate of *Other* weapons per group	Ordinal (0-4)
Group_Recordedtransfers	Number of recorded transfers of weapons per group	metric
Group_Recordedtransfers_SA	Number of recorded transfers of *small arms* per group	metric
Group_Recordedtransfers_LW	Number of recorded transfers of *light weapons* per group	metric
Group_Recordedtransfers_MCW	Number of recorded transfers of *major conventional weapons* per group	metric
Group_Recordedtransfers_Expl	Number of recorded transfers of *explosives* per group	metric
Transfer_Potential_Origin	List of potential origins of the armament transfers	String variable
Group_Evidenceentries	Number of entries with recorded evidence per group	Metric
Group_Evidenceentries_qnty	Number of entries with specific amounts given per group	Metric

Table 3: List of Variables Included in the Group-dataset (at the Lowest Level of Aggregation)

Variable name	Content	Scale
Group_SA_Machineguns	*Machine guns* armament per group	Ordinal (0-4)
Group_SA_Pistols_revolvers	*Pistols and revolvers* armament per group	Ordinal (0-4)
Group_SA_Rifles_Shotguns	*Rifles and shotguns* armament per group	Ordinal (0-4)
Group_LW_Heavymachguns_Cannons	*Heavy machine guns and cannons* armament per group	Ordinal (0-4)
Group_LW_Launchers	*Missile, rocket and grenade launchers* armament per group	Ordinal (0-4)
Group_LW_Mortars	*Mortars* armament per group	Ordinal (0-4)
Group_MCW_Aircraft	*Aircrafts* armament per group	Ordinal (0-4)
Group_MCW_Airdefence	*Air-defence systems* armament per group	Ordinal (0-4)
Group_MCW_Armouredvehicles	*Armoured vehicles* armament per group	Ordinal (0-4)
Group_MCW_Artillery	*Artillery* armament per group	Ordinal (0-4)
Group_MCW_Ships	*Ships* armament per group	Ordinal (0-4)
Group_MCW_Tanks	*Tanks* armament per group	Ordinal (0-4)
Group_Expl_Devices_materials	*Explosive devices and materials* armament per group	Ordinal (0-4)
Group_Expl_Landmines	*Landmines* armament per group	Ordinal (0-4)

4 Variables in the Groupyear-dataset

Table 4 describes the RAD-variables at the group-year level. Note that we only include ordinal measures in this version of the dataset since metric variables would imply a level of temporal precision that is not possible with yearly observations. At the group-year level, we also only cover the highest and medium level of aggregation.

Table 4: List of Variables Included in the Groupyear-dataset

Variable name	Content	Scale
Groupyear_Smallarms	*Small arms* armament per group-year	Ordinal (0-4)
Groupyear_Lightweapons	*Light weapons* armament per group-year	Ordinal (0-4)
Groupyear_MCW	*Major conventional weapons* armament per group-year	Ordinal (0-4)
Groupyear_Expolosives	*Explosives* armament per group-year	Ordinal (0-4)
Groupyear_Other	*Other* weapons per group-year	Ordinal (0-4)
Groupyear_Totalarms	Total armament per group-year	Ordinal (0-4)
Groupyear_SA_lowerbound	lower bound estimate of *Small arms* armament per group-year	Ordinal (0-4)
Groupyear_LW_lowerbound	lower bound estimate of *Light weapons* armament per group-year	Ordinal (0-4)
Groupyear_MCW_lowerbound	lower bound estimate of *Major conventional weapons* armament per group-year	Ordinal (0-4)
Groupyear_Expl_lowerbound	lower bound estimate of *Explosives* armament per group-year	Ordinal (0-4)
Groupyear_Other_lowerbound	lower bound estimate of *Other* weapons per group-year	Ordinal (0-4)

Group_Totalarms_lowerbound	lower bound estimate of the total armament per group-year	Ordinal (0-4)
Groupyear_Evidenceentries	Number of entries with recorded evidence per group-year	Metric
Groupyear_Evidenceentries_qnty	Number of entries with specific amounts given per group-year	Metric

5 Details on aggregating armament evidence

Table 5 displays the conversion of the metric variables into ordinal measures relating to the order of magnitude of the variables.

Table 5: Conversion between metric and ordinal variables

Value of the metric variable	Value of the corresponding ordinal variable
0	0
1-9	1
10-99	2
100-999	3
> 1,000	4

For entries where imprecise amounts are given, table 6 lists examples of how amount information provided in text sources were coded into ordinal values. The coding relates the verbatim given information to the implied order of magnitude, i.e. the value of the ordinal variables. Additionally, we also include entries just giving information on ammunition of specific arms, with ranking down these entries by one or two orders (for MCW and SALW, respectively) of magnitudes than entries stating information on weapons.

Coding imprecise amounts in this way allows to add up both precisely and imprecisely given amounts for each group-/group-year-category, resulting in the values for the ordinal lower bound variables. To be added up, we converted the ordinal values into numerical values, calculated the sum of the numerical values of both types of entries (with precise/numerical and imprecise/converted ordinal amount information given) and then re-converted the sums to ordinal values according to table 5. For the necessary conversion of ordinal values to numerical values, we calculated the mean values of entries *with precise* amount information in the respective value range, i.e. 2–9 for ordinal value 1[3], 10-99 for 2, 100–999 for 3, and the median of entries with > 1000 arms recorded for 4.

The other ordinal variables in the dataset represent our estimations additionally entries with missing amounts. For this, we assume that the amounts correspond to the

[3]Note that the range starts at 2, as this can be described as e.g. "multiple", whereas an amount of 1 will always be given in a precise term and can be recorded as numerical amount.

Table 6: Coding examples of imprecise amounts to ordinal values

Ordinal value	Source amount information
1	several, multiple, handful
	small/limited number
	few, some, a couple
2	dozens
	many, numerous
	1 truck (of light weapons)
	1 crate/box of Improvised Explosive Devices
3	hundreds
	1 truck (of small arms)
	large/great amount/numbers/cache
	tons (of small arms)
4	thousands
	trucks/truck loads (of small arms)

mean/median of other entries in the same group and category or, when the group does not have other entries coded in the same category, the same category in all groups, respectively. Be aware that this is a strong assumption that will be incorrect for each entry from an isolated perspective, but will allow for a more complete picture of rebels' armament in the overall perspective. Therefore, we provide the lower bound variables that give more confidence in each individual value, although missing out on relevant information for additional armament.

6 Additional details on creating the group-year version of the dataset

Attributing information on armament to an exact time span is not possible. For many entries, information on the specific point in time a weapon was in possession or was seized, for instance, is given. However, this does not include information on how long before or after, resp., this event the weapons were in possession of the group – it could be a few

days, the full year or decades. Adding up only the entries where the information stems from a given year to derive the value for this year would thus provide a very fragmentary picture of a group's armament, as most weapons will be in possession longer than would be indicated by the dataset.

To provide a more complete overall picture, we therefore prolonged the attributed time period for an entry. We create a 5-year shifting time window around each entry, employing the assumption that 5 years is a reasonable supposition for the duration weaponry is held on average. By employing these time windows to each entry, we derive a smoothed time series of overall armament, even though each value in a given year might be, in an isolated perspective, not be the "true" value. However, the alternative would be to miss out on large portions of armament that we have evidence for, but are not able to *exactly* attribute a valid time span to. Therefore, we propose this solution. Note that this, in our opinion, is also superior to providing the data in 5-year brackets, as this creates hard cut-off points regardless of the underlying recorded evidence, as well as to smoothing over the time series as a whole, as this would disregard *available* information on the time span for each entry.

For each entry, we know what type of evidence is provided, e.g. because of a seizure or because of leakage from state arsenals. Thus, we know that the relevant time period differs from entry to entry: when the evidence counts arms that are seized from a group, we know that the arms were in the group's possession *until* the event, i.e. the seizure, occurred. When the evidence counts arms that were leaked from state arsenals, we know that the arms were in the group's possession *after* the event, i.e. the leakage, occurred. Therefore, the 5 year window for seizures spans from $t-4$ to t, while the window for state leakage spans from t to $t+4$. Table 7 provides the respective time window for each type of evidence.

Note that we restrict the expansion of the time span to years within the active period for this group, defined as the period between the first and the last conflict-year recorded by UCDP within the 1989–2018 time window we restrict our data collection to.[4] For instance, when a group has its first conflict started in 2005, a seizure entry from 2007 only enters the aggregation for 2005, 2006 and 2007, as it is implausible for the group to

[4]Note that we do not restrict to active conflict-years, as complete disarmament between conflict periods is implausible.

Table 7: Valid 5-year time window for each evidence type

Type of evidence	$t-4$	$t-3$	$t-2$	$t-1$	t	$t+1$	$t+2$	$t+3$	$t+4$
Disarmament process	X	X	X	X	X				
Seizure by authorities	X	X	X	X	X				
Use in battle observed			X	X	X	X	X		
Evidence of possession			X	X	X	X	X		
General statement			X	X	X	X	X		
Battlefield capture					X	X	X	X	X
Loss after state collapse					X	X	X	X	X
Stockpile leakage					X	X	X	X	X
Transfer					X	X	X	X	X

possess these arms before the first conflict started. However, whenever evidence is found in a year outside the active period, we use it nevertheless, as it indicates actual possession in these years even though no conflict was recorded during this time.

7 Detailed descriptions and examples of weapon categories

This chapter provides definitions of the different weapon categories and their divisions into subcategories of weapon types as well as descriptions of these subcategories. The associations and the hierarchical organization of these categories are displayed in figure 1. These definitions build on the weapons classifications systems from SIPRI (2022), PRIO (Marsh et al., 2017) and NISAT (2007).

7.1 Major conventional weapons

With regards to *major conventional weapons* (MCWs), we follow the influential definition put forward by the Stockholm International Peace Research Institute which defines what they term "major weapons" as weapons systems that belong to one of several weapons categories (SIPRI, 2022).

In the Rebels' Armament Dataset, this includes the subcategories of *air-defence systems*,

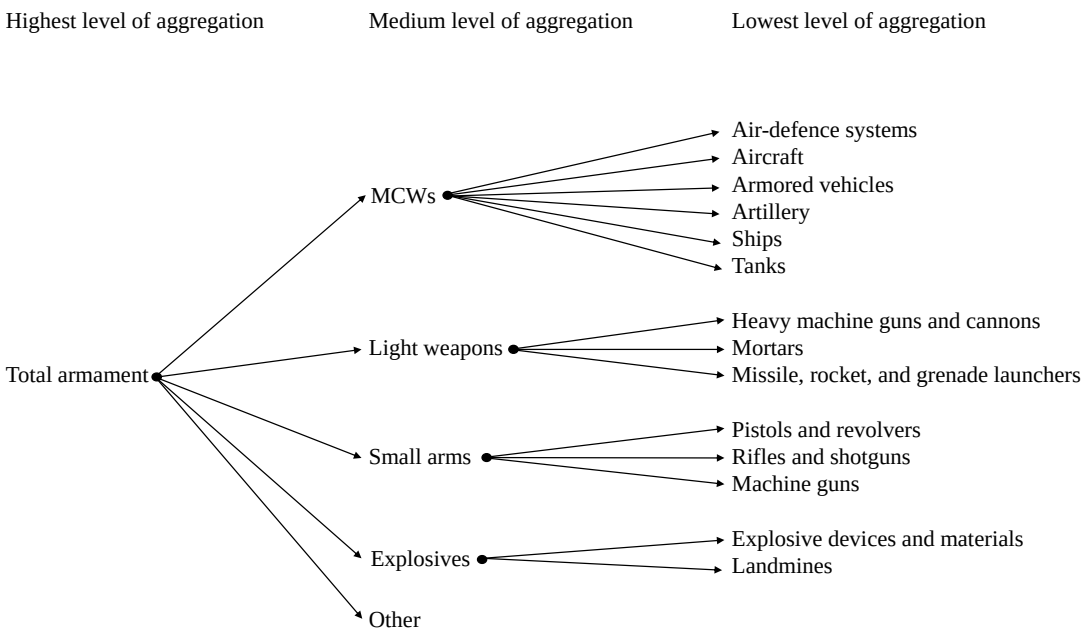

Figure 1: Depiction of the three different levels of aggregation and the hierarchical organization of (sub-)categories

aircrafts, *armored vehicles*, *artillery*, *ships*, and *tanks*.

7.1.1 Aircraft

The subcategory *aircraft* consists of (combat as well as transport) fixed-wing aircraft, (combat and transport) helicopters, and unmanned aerial vehicles (UAVs).[5]

7.1.2 Air-defence systems

Air-defence systems include all non-portable anti-aircraft cannons and surface-to-air missile systems.

7.1.3 Armored vehicles

Armored vehicles are vehicles with with integral armour protection. In practice, this includes mostly armoured personnel carriers, infantry fighting vehicles, and armoured trucks.

[5]In UN terminology, our *aircraft* category combines the categories "Combat aircraft and unmanned combat aerial vehicles (UAV)" and "attack helicopters" (United Nations, 2021) and additionally also includes transport aircraft.

7.1.4 Artillery

Artillery includes several non-portable weapon systems designed for long-range usage, such as howitzers, towed guns, and multiple (surface-to-surface) rocket launchers.

7.1.5 Ships

In the context of the armament of rebel groups, *ships* (which also includes boats) range from smaller naval vessels, such as armed speed boats, to larger vessels, e.g. gunboats.

7.1.6 Tanks

Tanks can be defined as "tracked or wheeled self-propelled armoured fighting vehicles with high cross-country mobility and a high-level of self-protection, weighing 16.5 metric tons unladen weight, with a high muzzle velocity direct fire main gun of at least 75 millimetres calibre" (United Nations, 2021). In contrast to SIPRI and in line with UN-definitions, we distinguish between *armoured vehicles* and *tanks* by the calibre of their main gun.

7.2 Light weapons

Light weapons (LW) are, "broadly speaking, weapons designed for use by two or three persons serving as a crew, although some may be carried and used by a single person" (United Nations, 2016).
The division of *light weapons* into subcategories closely follows the typification of PRIO (see e.g. Marsh et al., 2017). *Heavy machine guns and cannons*, *mortars*, as well as *missile, rocket and grenade launchers* are categorized as *light weapons* in the RAD.

7.2.1 Heavy machine guns and cannons

This subcategory includes a) heavy machine guns that are too heavy to be operated by a single person and are thus usually mounted e.g. on a vehicle and b) cannons, delimited by PRIO with calibres larger 12.7 mm.

7.2.2 Missile, Rocket and Grenade Launchers

Based on the UN-definition (United Nations, 2021), this subcategory includes portable weaponry systems that aim at delivering a warhead or weapon at targets on the surface or

in the air. Examples are rocket-propelled grenades (RPG) launchers, shoulder-launched anti-tank missiles, recoilless rifles, or Man-Portable Air-Defense Systems (MANPADS).

7.2.3 Mortars

Mortars are man-portable weapons that fire shells from a relatively short tube with high-arching trajectories.

7.3 Small arms

Small arms (SA) can be defined as, "broadly speaking, weapons designed for individual use" (United Nations, 2016).
We follow again the categorizations from PRIO and NISAT (2007). Here, *small arms* include *pistols and revolvers*, *rifles and shotguns*, and *machine guns*.

7.3.1 Machine guns

General-purpose machine guns, sub-machine guns, and machine pistols are included in this subcategory.

7.3.2 Pistols and Revolvers

Pistols and revolvers include different kinds of handguns, such as non-automatic, semi-automatic, automatic pistols, and revolvers.

7.3.3 Rifles and Shotguns

This subcategory is comprised of all non-automatic, semi-automatic, and automatic rifles, as well as shotguns of any kind (including, inter alia, assault rifles, sniper rifles, carbines, and pump-action shotguns).

7.4 Explosives

Landmines as well as *Explosive devices and materials* are contained in this category.

7.4.1 Explosive devices and materials

The subcategory *explosives* counts explosive material, improvised explosive devices (IEDs), explosive belts, bomb-making material, bombs, and (hand)grenades.

7.4.2 Landmines

Landmines include anti-personnel as well as anti-vehicle mines.

7.5 Other

This category captures all types of armament that do not belong in the categories *major conventional weapons*, *light weapons*, *small arms*, and *explosives*. In practice, it covers a considerable variety of entries, such as radar systems, communication equipment, electronic warfare systems, night vision devices, tear gas, smoke grenades, and non-armoured vehicles.

7.6 Additional information on arms types coding

Not all information found in public sources is consistently describing the specific type of weapon. In cases where given pieces of information are contradictory with regards to the categorization of weapons as *major conventional weapons*, *light weapons*, *small arms* or *explosives*, the following prioritization is applied. If available, the precise model name is considered the most relevant piece of evidence and used for the classification, followed by the general description (e. g. “rifle“, “cannon“ or “heavy machine gun“), and lastly (if applicable) the calibre of the weapon. In unclear cases, we also resorted to the existing NISAT (2007) and SIPRI (2022) armament classifications for comparable entries.

References

Shawn Davies, Therese Pettersson, and Magnus Ödberg. Organized violence 1989-2021 and drone warfare. *Journal of Peace Research*, 59(4), 2022.

Nils Petter Gleditsch, Peter Wallebsteen, Mikael Eriksson, Margareta Sollenberg, and Håvard Strand. Armed conflict 1946-2001: A new dataset. *Journal of Peace Research*, 39(5):615–637, 2002.

Nicholas J. Marsh, Topher L. McDougal, Ishtiaq Ahmed Khan, and Jorge Lison. Illicit small arms prices—transactions dataset: Codebook. pages 1–16, 2017.

Andreas Mehltretter, Oliver Pamp, Thurner Paul W., and Paul Binder. Introducing the Rebels Armament Dataset (RAD): Collecting Evidence on Rebel Military Capabilities. *Working paper*, 2023. URL `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4537283`.

NISAT. Small Arms Trade Database, 2007. URL `https://nisat.prio.org/Trade-Database/`.

Therese Pettersson. UCDP Actor Dataset Codebook v 22.1, 2022. URL `https://ucdp.uu.se/downloads/`.

SIPRI. SIPRI Arms Transfers Database - Methodology, 2022. URL `https://www.sipri.org/databases/armstransfers/background`.

United Nations. United Nations Small Arms Review Conference, 2016.

United Nations. United Nations Register of Conventional Arms. Categories of major conventional weapons, 2021. URL `https://www.unroca.org/categories`.