



MCARI

MYANMAR CONFLICT ANALYSIS
RESEARCH INSTITUTE

VERSION 1.0

MCARI Methodology and Codebook

The rules by which a report of an event becomes a row of data.

7

Event types

36

Sub-event types

37

Valid pairs

MCARI.ORG

Contents

	Introduction	3
1	The 24 published columns	5
2	The seven event types and their sub-events	8
3	The eight numeric fields	22
4	The three fatality buckets	29
5	Five coding rules that turn a note into a number	31
6	Worked examples	34
7	Geography	36
8	Actors	39
9	Limitations	43
Appendix A	The taxonomy at a glance	45
Appendix B	The standing wordings	47

Introduction

About MCARI

The Myanmar Conflict Analysis Research Institute (MCARI) records political violence and conflict-related events in Myanmar, event by event, from the day of the 1 February 2021 coup onward. One recorded event is one row of data. Each row states when and where the event happened, who acted and against whom, what kind of event it was, and how many people were killed, detained, displaced, released, or left one side for another.

MCARI's record is built from open-source reporting and from its own network of local sources. Reporting published by Myanmar outlets is monitored continuously.

Where several outlets report the same event independently, their accounts are read together and the agreement between them is the check. Where an event rests on one or two reports alone, it is put to people in the place concerned — residents of that township, or correspondents working there — and what they establish is recorded with the event. An event that local sources cannot confirm does not enter the dataset.

A claim made by a party to the conflict about its own operations is treated as a single-sourced report and checked the same way. Where sources disagree, the disagreement is recorded rather than resolved away.

What this document is

This document sets out the rules by which that record is made: what each published field holds, how the seven event types and thirty-six sub-event types are bounded against one another, what the numeric fields count and what they do not, how a quantity stated in words becomes a figure, how a place becomes a coordinate, and how a group's name is standardised.

Where a coding decision is not obvious, the reasoning behind it is given. A rule whose reason is unstated is a rule the next person will argue with, and a rule argued with is a rule applied unevenly.

It governs the dataset as a whole rather than any one period or any one report.

Who it is for

Three readers are expected, and the document is written for all three at once.

Coders apply these rules to produce rows, and need them stated precisely enough that two people reading the same report reach the same row.

Researchers and analysts working with MCARI data need to know what a field means before they count with it — what a zero is, what a blank is, what the fatality total does and does not include.

Reviewers assessing whether the figures can be relied on need to see the rules, the reasoning behind them, and the limits MCARI places on its own data. Those limits are set out in section 9.

What is published, and how to obtain the data

MCARI's quarterly and annual reports are published in full, in Burmese and in English. **The underlying dataset is not published in full. It is supplied on request**, and a request should state the period, the geographic scope, and the purpose it is needed for.

Two of the fields described in this document are never supplied. The narrative note is one of them, and the reason is not editorial: notes record what MCARI's local sources and correspondents established about an event, and releasing them would expose the people who did that work. The note is described here because it is the record every other field is derived from — not because it is available. The coding record that travels with each row is withheld for the same reason.

Versions

This codebook is versioned independently of the reports that cite it. A change takes a new version, and each version states the first issue it binds.

A report is never re-made against a later version. Each issue records the version its figures were coded under, so a reader can always tell which rules produced a given figure.

How to cite

MCARI, *MCARI Methodology and Codebook*, version 1.0 (U.S.: Myanmar Conflict Analysis Research Institute).

Data supplied by MCARI should be cited by edition:

MCARI, *Myanmar Conflict Data*, [edition], Myanmar Conflict Analysis Research Institute, supplied [date].

Contact

Myanmar Conflict Analysis Research Institute · mcari.org

1 The 24 published columns

The file carries 24 columns. Where a column's rules are set out elsewhere in this document, this section says what the column holds and points there rather than restating them.

Identity and time

mcari_id — the unique identifier for a row, stable across editions: an id is never reissued or renumbered, and deletions leave gaps in the sequence rather than closing them.

Two rows can agree on every published field and still be two events. Where one report describes several acts — protests by separately named groups, several villages shelled — each act is its own row. They are not duplicates, and the row count records acts.

event_date — the date the event occurred, never the date it was reported. ISO 8601, yyyy-mm-dd. MCARI never invents a date it was not given. Where a source describes a period rather than a day, the event is dated by what the source supports, and MCARI's working record notes how exact the date is: dated to the day, dated within a week, or dated more loosely.

Per-day dates, or a start plus a completed duration ("over the past five days"), give one row per day, counted back from the reporting date, each dated to the day; expansion is bounded at one month, and a longer span takes a single row. An approximate duration ("roughly five days") cannot be counted back to exact days and gives one row at the midpoint, dated more loosely. A count with no reconstructable dates ("47 clashes over ten days") gives one row at the midpoint of the stated span, dated within a week where the span is 7 days or fewer and dated more loosely where it is 8 days or more; the event count then undercounts, which is the cost of not guessing. A start day with no end gives one row on the start day, dated more loosely.

An event with no day named that the source presents as current — "in recent days", "this week" — and that carries an actor, a place and an act, or a figure, gives one row at the publication date, dated more loosely. It is the only case in which the publication date becomes **event_date**. An old event the source cites as the past with no day and no month is a background reference and is not recorded, however complete the rest; nor is an event that is only referred to, without an actor, a place and an act.

Classification

event_type — one of seven values. Defined in section 2.

sub_event_type — one of 36 strings, forming 37 valid pairs with **event_type**. Listed and defined in section 2. A validator checks the pair, not the string alone.

Actors

actor1 — who carried the act out. **actor2** — the primary target or opposing force. **assoc_actor_1** and **assoc_actor_2** — others on the first and second actor's side.

Names are standardised against MCARI's actor registry rather than being taken from a source's own wording, so that one group is one string across the series. An unidentified perpetrator is still

an actor and is recorded as one. A blank actor₂ on a Sabotage row is correct rather than missing: the act has a target that is not a party.

Geography

admin1 state or region · **admin2** district · **admin3** township · **pcode_admin3** the township's MIMU P-code · **latitude** and **longitude** the event's point location · **geo_precision** the administrative level admin3 actually reached. All seven are set out in section 7.

Source

source — the outlet or outlets the event was reported by. Where more than one reported it, the names are joined with ; . The outlets are not classified by type, so the share of events resting on a conflict party's own account cannot be separated out from the data as supplied.

Quantities

fatalities · **fatalities_military_soldiers** · **fatalities_non_military_armed_groups** · **fatalities_civilians** · **arrested** · **displaced_count** · **defected_count** · **released_count** — the eight numeric fields, set out in section 3, with the three fatality buckets in section 4. Every field on every row carries a value.

The coding record

Beside the note, each row carries a short coding record in a field named `conflict_flag`. It holds what a reader of the row would otherwise have to guess: which of two readings was taken and why, a figure a source gave that was not the one recorded, the wording a source used for a quantity that had no number, a discrepancy between outlets. It is referred to throughout this document, because most of the rules below end by saying what is written there.

Like the note, it is not supplied with the data. It is described here because it is where the reasoning behind a published figure is kept.

The note

The note is not a column of the published file, and it is the method. Every classification field in this document is derived from it.

A note is a single narrative account of a single event, compiled at the stage between reading the reporting and coding it. One event produces one note, whatever the source material: where several outlets reported the same event, their accounts are read together and the note records what they jointly establish, with disagreement between them carried rather than resolved away. Where one report describes several distinct events, it produces several notes, and each becomes its own row.

The note has a fixed form — on [date], in [place], [actor] [action], [casualty status], [any charge, release or context] — and casualty status always closes it explicitly, so whether anyone was

harmed is never left to be inferred from silence. Notes are written in neutral English, and loaded terminology from a source does not enter one, even in quotation marks. Every group a source names appears in the note; none is compressed away, because a group dropped from a note is a relationship between actors lost permanently.

The note is the primary record and the single source of truth for the row. Classification fields are derived from it rather than from the original reporting, which is what keeps a field from drifting out of agreement with the account it rests on, and what allows a change in the taxonomy to be applied by re-reading the notes rather than the articles. Where one report produced several rows, those rows carry the same note text — see section 3 for what that means for a zero on one of them.

2 The seven event types and their sub-events

Shelling / Artillery / Missile attack sits under two types, so 36 strings form 37 pairs. Both counts are correct — they measure different things. Validators must check the pair, not the string alone. Where a sub-event is described with reference to international law, the reference explains why the category is recorded; MCARI records what its sources report and makes no legal determination about any event, party or person.

Event type	Sub-event types
Armed Conflict	8
Violence against Civilians	5
Targeted Violence	1
Coercion	7
Protests / Civil Unrest	2
Strategic Developments	13
Other	1
Pairs	37

Armed Conflict

Definition. Direct military engagements between armed actors.

SUB-EVENT TYPES

Armed clash

Direct fighting between two or more armed groups involving exchange of fire or physical combat. Includes ambushes, raids, and offensive operations where both sides are armed combatants. Also the default where a place was attacked and no method sub-event fits — arson, burning, a post set alight after being abandoned — with the method recorded alongside the row. There is no arson sub-event under Armed Conflict, and Arson / Property destruction under Violence against Civilians is the civilian-property bucket only.

Airstrike

Attack carried out by military aircraft (fixed-wing jets or helicopters) using bombs, rockets, or guns. No drone involvement. For a strike that is the event itself, with no clash under way.

Drone attack

Attack carried out by an unmanned aerial vehicle (UAV/drone), including commercial drones modified for weapons delivery. Must be confirmed in the source.

Shelling / Artillery / Missile attack

Attack using artillery, mortars, rocket launchers, or missiles fired from ground-based systems. Use when an opposing armed actor is named as the target, or when fire is aimed at believed enemy presence. The belief need not be correct — what matters is that the note states the fire was aimed at something.

Landmine / IED

Detonation of a landmine or improvised explosive device that was emplaced as an attack. Covers all victims — combatant or civilian — because the device was placed to attack. Use *fatalities_civilians* for civilian deaths. A mine detonation followed by return fire is one row here; troops dismounting and exchanging fire does not create a second event.

Grenade attack

A thrown or launched explosive (grenade, pipe bomb, homemade bomb) fired directly — a thrown grenade, or an RPG or shoulder-fired launcher aimed at a target in the firer's line of sight. Indirect fire (mortar, artillery, rocket-artillery) is *Shelling / Artillery / Missile attack*; where the note does not say whether fire was direct or indirect, the existing sub-event stands. A grenade aimed at property with no one struck — a parked vehicle, an office, a shopfront — stays here.

Sabotage

Deliberate destruction of inert, unoccupied infrastructure outside active combat — a bridge, a pylon, a rail line, a telecommunications tower. Applies symmetrically: resistance destroying SAC infrastructure, and SAC destroying resistance infrastructure, are both Sabotage.

Building type decides, not who happened to be inside at the time. A building in use — a shop, a bank, a company or administrative office, a government building, a barracks or camp — is personnel-occupied whatever the hour, and is never Sabotage: it is classified by method. Where someone is struck at an inert target, the row stays here and the fatality fields record it.

Siege / Blockade

Military forces surround and cut off access to a town, area, or location, preventing movement of people, supplies, or reinforcements in or out.

WHERE THE BOUNDARY FALLS

What decides the category

Governs every type

- Core principle: target first, method second.
- Friendly fire is coded by the act: a strike that also hits the firing side's own forces keeps the act's classification, with the own-side deaths in their bucket on that row. Other / Other is only for an accident that is the whole event — an own device, unexploded ordnance.

Was a two-sided engagement already under way, or was the strike the event itself?

Armed clash ↔ Airstrike · Drone attack · a method sub-event

- A one-sided attack on a position takes a method sub-event only where the note states the method. Where no weapon is named → Armed clash.
- An airstrike or barrage delivered during an active two-sided engagement is part of that engagement — the row stays Armed clash, with the air support in *conflict_flag*.
- A drone strike met with return fire stays here — the strike is the act that defines the incident. Only where the note makes a prolonged battle its main subject → Armed clash.

Aircraft or drone?

Airstrike ↔ Drone attack

- If a drone is confirmed → Drone attack

Fire from a distance: was an opposing combatant the target, or a civilian place?

Armed Conflict ↔ Violence against Civilians / Shelling / Artillery / Missile attack

- Where the note names an opposing combatant as the target, or says fire was aimed at believed enemy presence → Armed Conflict / Shelling / Artillery / Missile attack
- Where the note says fire was aimed at believed enemy presence, the row stays Armed Conflict even when civilians are killed, and their deaths are recorded in `fatalities_civilians`. A civilian killed or struck decides the row only where the note names nothing as the target.
- "Fired towards the village" is fire directed at a settlement → this row, casualties known or not.
- The preposition is never the test. "In X village" opening a note is the location preamble, not the target.
- Unpopulated-ground floor — fire that struck neither an opposing force nor a civilian target (a hillside, a riverbank, open country), or a note naming nothing struck at all, stays here.
- A displacement figure with nothing named as struck decides nothing — the row stays here and the figure stays on it.

An explosive: was it aimed at people, at a structure, or at civilian property?

Grenade attack ↔ Sabotage ↔ Violence against Civilians / Arson / Property destruction

- Intent and target decide, not method. Where the target is inert, unoccupied infrastructure — a bridge, a pylon, a rail line, a tower → Sabotage, whatever was used. A building in use is not that, and takes the method instead.
- A personnel-occupied post — station, outpost, barracks, checkpoint, gate — is never Sabotage, even where only the structure was damaged; it is classified by method.
- Artillery, mortar or rocket fire is never Sabotage whatever the target → Shelling / Artillery / Missile attack
- If civilian property (homes, markets, religious buildings) is destroyed → Violence against Civilians / Arson / Property destruction
- A grenade, mine or IED aimed at property is an explosive attack, not arson → Armed Conflict, by method

A vehicle attacked with an explosive?

Sabotage ↔ Grenade attack · Landmine / IED

- A grenade, mine or IED aimed at a parked vehicle, no one struck, stays Armed Conflict, by method — a vehicle is not inert infrastructure. A military or police vehicle is a mobile position (section 5, rule 1); a mine or IED initiating the attack is Landmine / IED whoever is inside.

A device that exploded: was it emplaced to attack?

Landmine / IED ↔ Other / Other

- A device emplaced to attack stays Armed Conflict / Landmine / IED whoever the victim. Only a device that detonates without an attack comes here.

- Where the note does not say whether the device was emplaced or abandoned, nothing is guessed — the row is not moved to Other / Other, and the uncertainty is recorded with it.

Cut off, or restricted?

Siege / Blockade ↔ Coercion / Intimidation ↔ Strategic Developments / Political statement / Declaration

- An armed force cutting off a place or an opposing force as a military measure — encircling a town or base, closing roads, rivers or bridges to stop supply reaching ground the other side holds or contests → Siege / Blockade.
- Checkpoints, curfews or bans on entering or leaving villages that a force enforces on civilians in ground it controls → Coercion / Intimidation.
- An announced order with no enforcement described → Strategic Developments / Political statement / Declaration.

Set out under another type: Capture during a clash — under Coercion · A person or a place — under Targeted Violence · Holding a captured place — under Strategic Developments · Surrender or capture; weapons taken in a clash — under Strategic Developments.

Violence against Civilians

Definition. Deliberate violence targeting non-combatants.

SUB-EVENT TYPES

Attack on civilians

Physical attack targeting civilian persons or groups using ground forces, small arms, or close violence. Applies to indiscriminate or group-level attacks where the target is the civilian population rather than a specific named individual.

Sexual violence

Sexual assault, rape, or other forms of conflict-related sexual violence perpetrated by armed actors against civilians.

Arson / Property destruction

Deliberate burning or destruction of civilian property including homes, villages, markets, and religious buildings. The target must be civilian property, not military or government infrastructure. This is the only civilian-property bucket; it carries no fatality.

Shelling / Artillery / Missile attack

Indirect fire or air-delivered munitions striking a civilian target — a village, a gathering, a ceremony, a residential area — with no opposing combatant named as the target. A combatant among the casualties does not move the row: what the note says was struck decides it, not who was hit. This sub-event is dual-homed: it covers airstrikes and drone-dropped bombs on civilian targets, because Violence against Civilians has no separate Airstrike or Drone attack sub-event.

Mob violence

Violence perpetrated by a civilian group or mob against other civilians. The perpetrator is not an organised armed group but a civilian crowd, spontaneous or organised. actor1 is Rioters.

Close violence, or fire from a distance?

Attack on civilians ↔ Shelling / Artillery / Missile attack

- For indirect fire or air-delivered munitions on civilian targets → Violence against Civilians / Shelling / Artillery / Missile attack

A death in custody: one row or two, and which?

Attack on civilians ↔ Coercion / Arrests / Detention

- A detainee who dies in custody never opens a second row. The death is recorded in the fatality field of the single row that already exists for the detention.
- Detainees executed — the purpose was to kill → that single row is Attack on civilians, with arrested and fatalities both on it — except a member of an armed formation, which is Targeted Violence / Targeted killing (under Targeted Violence); one shot while escaping stays here.
- Detainees who die from beating or mistreatment → that single row stays Coercion / Arrests / Detention, with arrested and fatalities both on it.
- Only where the arrest and the death fall on different dates does this become two rows, one on each date.

A death in a fire, or a sexual assault that ended in a death?

Attack on civilians ↔ Arson / Property destruction · Sexual violence ↔ Attack on civilians

- Deaths in a fire an actor set are one act — the row is Attack on civilians with the fatality on it, not a separate Arson / Property destruction row.
- Where the same victim was also killed, this is one row coded here, with the fatality recorded on it. The more specific sub-event takes precedence, and splitting would double-count one person.

Property taken, but nothing destroyed and nobody moved?

Arson / Property destruction ↔ Coercion / Forced displacement

- Property confiscated, locked or seized with no destruction and no forced relocation → this row

Set out under another type: The target of fire from a distance; an explosive aimed at property — under Armed Conflict · A person attacked for a role — under Targeted Violence.

Targeted Violence

Definition. Deliberate killing of specific individuals.

Qualifying roles — SAC side: Administrator (any level) · military informant · Pyu Saw Htee or any named pro-military militia · USDP member · Police · soldiers and officers of any rank · SAC-appointed official · BGF member

Qualifying roles — Opposition side: NLD officials, candidates and election staff · pre-coup government officials · CRPH and NUG-affiliated figures · CDM leaders and organisers · PDF and EAO members

SUB-EVENT TYPES

Targeted killing

Deliberate killing of a specific person, or an attempt made with lethal means, because of their role or affiliation in the post-coup conflict — on either side. The test is counterfactual: if that person had not been there, would this attack still have happened? The role must be a formal position (member, official, candidate, staff, organiser) stated in the note's own voice. A role stated only as an allegation — by the killers, by local people, on a note left with the body — does not qualify; where the note states the role in its own voice and alleges only what the person did or why they were killed, the role stands. A supporter or sympathiser with no formal role does not qualify. Retired and former officeholders default to Civilians unless the note shows an ongoing connection, or shows they were singled out for that former identity. Prominence is not a factor.

A fighter captured and then killed is recorded here. The victim is not a civilian, and the killing rests on who they were; killing a person already placed *hors de combat* is a named war crime under the Rome Statute, Art. 8(2)(c)(i). Where the capture and the killing fall on the same date they are one row. Where the killing comes later, the capture is recorded on its own date as Arrests / Detention and the killing on the date it happened. The coding record names the capture in both cases, so these killings can be found again.

A detainee shot during an escape attempt is not recorded here. Where the note gives an escape, the killing rests on the flight rather than on who the person was, and the row stays Violence against Civilians / Attack on civilians, or the clash row it belongs to. The paragraph above reaches only a killing the note presents as an execution.

WHERE THE BOUNDARY FALLS

A person attacked for a role, or civilians attacked as civilians?

Targeted killing ↔ Violence against Civilians / Attack on civilians

- Distinction from Attack on civilians: Targeted killing = a person attacked for a qualifying role. Attack on civilians = civilians attacked as civilians.
- A role stated only as an allegation, whoever makes it, is not a qualifying role. An allegation about what the person did, beside a role the note states in its own voice, leaves the role standing.

A person, or a place?

Targeted killing ↔ Armed Conflict

- A fighter killed while the fighting was still under way is part of that engagement — the clash row carries the death. One captured, and killed after the fighting had ended for him, is this row.
- A place or position attacked — post, checkpoint, booth, outpost, base — is Armed Conflict, even one-sided.
- A military or police vehicle is the mobile equivalent of a position → Armed Conflict. A private vehicle attacked because a qualifying person is inside → this row.

Lethal means, or beating and threats?

Targeted killing ↔ Attack on civilians · Arrests / Detention

- A qualifying-role person beaten or threatened, not attacked with lethal means → Violence against Civilians / Attack on civilians, or Coercion / Arrests / Detention where it happened in custody.

Both actors on the same side?

Targeted killing ↔ Other / Other

- Holds even where both actors are on the same side — a role or allegiance makes it this row; a personal dispute makes it Other / Other.

Coercion

Definition. State or armed group coercive control over individuals.

SUB-EVENT TYPES

Arrests / Detention

Arrest or detention of persons by security forces or armed groups, where the detention is the primary act — a house raid, a checkpoint stop, a targeted arrest. Applies to both political arrests and arrests of armed group members.

Enforced disappearance

A person is taken by an armed actor — the military, police or pro-military militia, or a resistance or ethnic armed group — and the note states that their fate or whereabouts are unknown, or that the holder denies or will not disclose holding them. MCARI codes the act, not the actor's status. The taking fills arrested.

Forced displacement

An actor drove people out — an order to leave, homes burned or destroyed, a forced relocation. It does not mean people fled in fear. Use only where the displacement is the primary reported event and the note describes no other codeable act.

Forced conscription / Porter duty

Forcible recruitment into military service, or compulsion to serve as a porter carrying military supplies or guiding a route. Records a taking — people actually seized. Record the number in arrested.

Human Shields

Civilians taken and used to protect an armed actor from danger — walked ahead of troops, placed in vehicles, or held at a position under threat of attack. Prohibited under customary international humanitarian law in international armed conflicts and, arguably, in non-international armed conflicts (ICRC, Customary IHL, Rule 97); the Rome Statute lists it as a war crime in international armed conflict, Art. 8(2)(b)(xxiii). Recorded once, as the act of taking them, at the place and date the note first records it.

Communications blackout

Deliberate disruption or shutdown of telecommunications, internet, or mobile networks by an armed actor or state authority, restricting civilian access to information and communication. Includes internet shutdowns, SIM card deactivations, and phone network blackouts imposed on a township, district, or region.

Intimidation

Coercive actions that create fear without direct physical harm or formal arrest. Includes raiding hospitals, schools, or offices; recording names of civilians; door-to-door threats; enforcing movement restrictions with troops; and surrounding residential areas. No arrest and no confirmed physical injury occurs.

WHERE THE BOUNDARY FALLS

Captured during a clash, or detained as an act of its own?

Armed Conflict / Armed clash ↔ Arrests / Detention

- Capture during an ongoing clash is an outcome of the clash → stays on the Armed clash row with arrested filled there.
- The line is timing, not status. Where the note places the capture after the person broke away — taken while fleeing, found in hiding, handed over later — the detention is its own act → this row.

What was done to the people held, and by whom?

Arrests / Detention ↔ Violence against Civilians

- People detained and beaten or abused in the same event → this row, arrested filled, fatalities 0, the abuse in conflict_flag.
- People already in custody from a separately recorded event and abused later — prisoners beaten in a prison — → Violence against Civilians / Attack on civilians, arrested 0; an arrest count is never carried forward.
- A non-state militia detaining a civilian is still this row — non-state status does not change it.

Held, or taken and hidden?

Arrests / Detention ↔ Enforced disappearance

- The note says, or makes plain, who holds the person or where — a police station, a base, a charge, a release → Arrests / Detention.
- A taking, and the fate or whereabouts unknown or the custody denied → Enforced disappearance. Both are required.
- A person who "went missing" with no taking stated is neither.

Taken and hidden, or taken and used as a shield?

Enforced disappearance ↔ Human Shields

- Not Enforced disappearance — people used as shields are usually released or killed, not disappeared.

Taken to carry, or taken to shield?

Forced conscription / Porter duty ↔ Human Shields

- Labour — carrying supplies, guiding a route → Forced conscription / Porter duty
- Both protection and labour → Human Shields (the graver violation)

Was anyone actually taken?

Forced conscription / Porter duty ↔ Intimidation ↔ Strategic Developments / Military training / Graduation

- A conscription list, a village quota, or a summons with nobody taken is pressure applied, not people seized → Coercion / Intimidation.
- Unwillingness does not make a training a taking. People actually seized and taken → Forced conscription / Porter duty, with arrested filled.

Did people flee an act, or were they driven out?

Forced displacement ↔ the act that caused the flight

- Where the note names the act that caused the flight — shelling, an airstrike, a raid — code that act and put the figure in `displaced_count` on that row. Do not open a second row here.
- A raid or an advance with no violent act named is not itself a codeable act → this row is correct.

A network shut down, or a tower destroyed?

Communications blackout ↔ Armed Conflict / Sabotage

- If the note describes physical destruction of towers or cables → Armed Conflict / Sabotage

An order, or troops enforcing it?

Intimidation ↔ Strategic Developments / Political statement / Declaration

- Movement restriction order (administrative, decree) → Strategic Developments / Political statement / Declaration
- Physical enforcement of a movement restriction by troops → this row

An administrator resigns: was anyone applying pressure?

Intimidation ↔ Strategic Developments / Defection / Surrender ↔ Other / Other

- An administrator threatened by a resistance group into resigning → this row, not Defection / Surrender
- An administrator who resigns with no actor named as pressuring them → Other / Other, `defected_count` 0.

Set out under another type: People compelled to demonstrate — under Protests / Civil Unrest · A death in custody — under Violence against Civilians.

Protests / Civil Unrest

Definition. Collective civilian action; political expression people chose to make.

SUB-EVENT TYPES

Peaceful protest

Organized demonstration, march, strike, or sit-in by civilians without security force intervention or violence. Records people assembling of their own will.

Protest with intervention

Protest during which security forces intervene to disperse demonstrators using force, tear gas, rubber bullets, or live ammunition. The violence is initiated by security forces, not protesters.

WHERE THE BOUNDARY FALLS

Chosen, or compelled?

Peaceful protest ↔ Coercion / Intimidation

- A turnout an armed actor compelled is not a protest → Coercion / Intimidation

One protest, or several?

One row ↔ several rows

- Separately named groups each holding their own protest are separate rows, in one township as much as in two.

What else happened at the protest?

Protests / Civil Unrest ↔ Coercion / Arrests / Detention · Violence against Civilians

- Protest + arrests → split (this row + Coercion / Arrests / Detention). Protest + fatalities → split (this row + Violence against Civilians / Attack on civilians). A protest row carries `Protesters` in `actor1`. Recording the arrests on it would put them in the field that says who acted — that is, it would record the protesters as having made the arrests. The split keeps each act with the party that carried it out.
- Protesters themselves burn or vandalise a government building → split (Peaceful protest, `actor1` Protesters + Violence against Civilians / Arson / Property destruction, `actor2` = the government body)

Strategic Developments

Definition. Politically or militarily significant non-violent events.

SUB-EVENT TYPES

Ceasefire / Peace talks

Formal or informal ceasefire agreements, peace negotiations, or diplomatic talks between armed parties. Includes both announced agreements and ongoing negotiation processes.

Territorial control change — Armed opposition

PDF, EAO, or resistance forces capture and hold a town, military base, or strategic location previously held by the Myanmar Military or SAC-aligned forces.

Territorial control change — Military

Myanmar Military recaptures and holds territory, towns, or bases previously lost to resistance forces.

Headquarters / Base established

An armed group formally establishes a new permanent or semi-permanent headquarters, military base, or camp. Refers to an established physical presence, not temporary deployment or territorial capture.

Political alliances

Formal announcement of a political or military alliance between two or more armed or political actors. Includes coalition agreements, joint operation declarations, and formal partnerships. Formations merging, reorganising into a joint command, or placing themselves under an existing structure belong here — that is an alliance change, not a new formation.

Political statement / Declaration

Official statements, announcements, or declarations by armed or political actors with strategic significance. Includes policy announcements, international sanctions, administrative orders, and calls to action. No active incident occurs.

Defection / Surrender

An armed actor leaving one side for another. The note must name the side received into — surrendered to, joined, defected to. A soldier or police officer who joins the Civil Disobedience Movement has been received by a side: the opposition calls on them to do so and supports those who do.

Prisoner release

Release by an actor of people it was holding — detainees, prisoners of war, or political prisoners. Includes negotiated releases, unilateral releases, amnesties, and prisoner exchanges.

Forces deployment / Positioning

Reported movement or positioning of armed forces for a confirmed period (days/weeks), without active combat or coercive action against civilians. Includes troop buildup reports, reinforcement movements, withdrawals and strategic repositioning.

Military clearance operation

Military forces enter a village or area with coercive intent — searching homes, detaining residents, or creating fear — without active combat occurring. Civilians typically flee.

New armed group / Faction formation

Formal announcement of the formation of a new armed group, resistance unit, or political-military faction. For a formation that did not exist before. Includes newly formed PDF units, EAO splinter groups, and pro-military militias.

Weapons seizure

Confiscation or capture of weapons, ammunition, or military equipment outside active combat. Includes checkpoint seizures, and house/monastery/base raids where weapons only are seized with no arrest and no violence.

Military training / Graduation

An armed group completing or holding military training, including graduation ceremonies. Recorded because training completion is a force-generation signal — a real indicator of a group's growth. Where the note records training held and nobody was seized, this is the correct row with arrested 0, and any unwillingness recorded in conflict_flag.

WHERE THE BOUNDARY FALLS

Captured and held, or only attacked?

Territorial control change — Armed opposition · Territorial control change — Military ↔ Armed Conflict / Armed clash

- This row opens only where the note says the place was captured AND held — captured, taken control of, seized, occupied, or regained, retook, recaptured for a place passing back to the side that held it before. It sits beside the Armed clash row, on the same date, not instead of it.
- Attacked, overrun, burned, raided, withdrawn from, abandoned is not holding — the clash row stands alone.
- Holding is never inferred from a successful attack. The note must use one of the holding words.
- A capture a side claims is recorded as a capture. Where the note says a side claimed to have captured a place, the row opens as it would for a stated capture, and the coding record says it rests on that side's claim.
- Two cases open no row: a capture the other side denies within the same report, and a place taken and lost again the same day.
- A road is not a place that can be held.

Troops arrived: for how long?

Headquarters / Base established ↔ Military clearance operation ↔ Forces deployment / Positioning

- Troops with a confirmed duration (days/weeks) → Forces deployment / Positioning

Troops entered: with violence, or without?

Military clearance operation ↔ the violent types

- Requires no violence. If looting, arson or shooting also occurs, classify by the violence instead.

Received by another side, or simply gone?

Defection / Surrender ↔ Forces deployment / Positioning

- A facilitator is not a receiving side. "Defected with the help of X" names who helped him leave, not who took him in.
- Joining the Civil Disobedience Movement is joining a side → this row.
- Deserting, abandoning a post, or fleeing across a border with no receiving side named → Forces deployment / Positioning, defected_count 0.
- Troops or police withdraw or retreat with no clash described → Forces deployment / Positioning, "retreat/withdrawal" in conflict_flag. A Territorial control change row opens on the same date only where the note says the other side then held, occupied or seized the place.

Given up, or taken?

Defection / Surrender ↔ Armed Conflict / Armed clash

- Laying down arms and giving themselves up to the force fighting them is a surrender, not a capture → this row. Where the note gives no number, the row still opens, with `defected_count` 3 and `Defection or surrender occurred`; number unknown.
- Being captured (taken, not given up) during a clash → stays on the Armed clash row with `arrested` filled. Read the verb the note uses.

Released by the holder, or freed some other way?

Prisoner release ↔ Other / Other

- Extraction or rescue by an outside party is not a release — the holding actor released nobody → `Other / Other`.
- A handover to another state's authorities is a transfer of custody, not a release → `Other / Other`.

A new formation, or existing ones combining?

New armed group / Faction formation ↔ Political alliances

- Formations merging, reorganising into a joint command, or placing themselves under an existing structure → `Political alliances`, not this row.

Weapons alone, or with a clash or an arrest?

Weapons seizure ↔ Armed Conflict / Armed clash · Coercion / Arrests / Detention

- Weapons taken as an outcome of a clash → no separate row; record in the clash row's `conflict_flag`
- If an arrest also occurs in the same raid → `Arrests / Detention` row, weapons in `conflict_flag`

| *Set out under another type:* Training and seizure — under `Coercion` · An administrator who resigns — under `Coercion`.

Other

Definition. Events that do not fit any existing category.

SUB-EVENT TYPES

Other

For an event the taxonomy genuinely has no home for — not for one that is merely unusual. In practice: harm with no attacker (leftover ordnance, UXO, a device detonating on its handlers, a bomb defusal); harm inside a single force or outside the conflict's two-sided structure (an internal killing with no opposing actor, a civilian who kills the person seizing him); and outcomes that resemble a recorded category but are not it (detainees extracted by a third party, or handed to another state — neither is a `Prisoner release`).

A killing inside one force?

Other ↔ Violence against Civilians

- An internal killing with no opposing actor is not Violence against Civilians — the victim was not a civilian.

Set out under another type: A device with no attack behind it — under Armed Conflict · A personal dispute within one force — under Targeted Violence · An administrator who resigns unpressured — under Coercion · Extraction or handover of detainees — under Strategic Developments.

3 The eight numeric fields

The eight are fatalities, fatalities_military_soldiers, fatalities_non_military_armed_groups, fatalities_civilians, arrested, displaced_count, defected_count and released_count.

Every field on every row carries a value

Blank is not a valid final state. Each field is decided on its own, by what the note says about that one quantity:

What the note says about that quantity	Value	conflict_flag
Nothing at all	0	nothing
It happened, no number given	3	[quantity] occurred; number unknown. — for fatalities, Unknown fatalities coded as 3.
The source does not say whether anyone was killed — fatalities only	0	Casualties unknown.
A number	that number	nothing

The cases are decided per field, not per row. Two of them are easy to run together, and they carry different values:

Standing wording	What the note did	fatalities
Casualties unknown.	the source does not say whether anyone died	0
Unknown fatalities coded as 3.	a death is stated with no count	3

Each wording resolves its own quantity and no other: either of the two above settles fatalities and says nothing about arrested, displaced_count, defected_count or released_count, each of which is read on its own. The three fatality buckets stay 0 where the note does not say which side the dead belonged to.

There is no injured field. The dataset has fields for five quantities — killed, arrested, displaced, surrendered, released. An injured figure is recorded in the note text and nowhere else.

Quantifying vague terms

Where a source gives a quantity in words rather than a figure, MCARI assigns the value below. Scale words take the minimum the wording confirms; 3 is a fixed convention for a small number the source does not give. A vague descriptive quantity indicates that an event occurred; therefore, it is assigned a specific baseline value rather than a zero.

Source's wording	Assigned figure
some · several · a few · a number of · many	3
dozens	20
hundreds	200
thousands	2,000
tens of thousands	20,000
hundreds of thousands	200,000
about N · some N · nearly N · almost N	N
more than N · over N · at least N	N
an event stated with no count (any field)	3

N is the figure the source itself states. some followed by a number — "some 15 boys" — means about that number and takes N, not 3.

The quantity words — dozens, hundreds, thousands, tens and hundreds of thousands — take the minimum the wording confirms.

3 is a convention applied consistently since 2021, not a minimum. Every row carrying it also carries the standing wording, so the convention figures are marked wherever they appear.

No subtraction. These derived figures are never used as the basis for subtraction. If a source states 60 people were detained and "over 50" were released, MCARI does not conclude that 10 remain in custody. 50 is only a minimum threshold, so the exact remaining number is unknown.

fatalities

What it counts. Deaths. MCARI counts deaths inflicted by one party on another.

A death from wounds. Where a person wounded in an attack dies of the injuries on a later date, the death is recorded on the row of the act that caused the wounds, and the date of death in the coding record. No row opens on the date of death: the death is not a new act. This differs from a detainee killed on a later day, which is a second act and takes its own row.

WHAT IT DOES NOT COUNT.

- **A self-inflicted death.** A suicide following a killing is recorded in `conflict_flag` and left out of fatalities. Counting it would inflate the conflict toll.
- **A death in custody that no party inflicted.** Illness, a seizure, or another cause the note attributes to no one's act — and a suicide in custody — is recorded in the coding record and left out of fatalities.
- **The wounded.** A combined killed-and-wounded figure is not a death count. The fatality fields hold deaths only, and the combined figure goes in `conflict_flag`.
- **A cumulative or organisational total.** An organisation's total over a period — the UN's, a rights group's, a monitoring group's, or the military's own cumulative claim — is not recorded, whatever the period's length.

- **A figure another row already carries.** See the two rules below on figures split and repeated across rows.

The three fatality buckets

What each bucket counts and does not count is set out in section 4. Two limits belong here: the three buckets never sum to more than their own row's fatalities, and they do not have to sum to it at all.

arrested

What it counts. People seized by an actor, whatever sub-event carries the act. A Human Shields headcount and the number taken in an Enforced disappearance both record people held against their will, and both fill arrested on their own row. Forced conscription / Porter duty fills it the same way: conscripts and porters are seized and held exactly as detainees are, and the count is the only place that quantity is recorded. The field is not reserved for Arrests / Detention.

WHAT IT DOES NOT COUNT.

- **Two figures at once.** Where a source reports a detention count alongside a shields headcount, the shields figure is the row's value and the detention figure stays in conflict_flag.
- **People not taken.** A conscription list, a village quota or a summons with nobody seized is pressure applied rather than people seized: the row is Coercion / Intimidation and takes arrested 0. So does a training nobody was seized at.
- **People already held from a separately recorded event.** An arrest count is never carried forward onto a later row.
- **Counts carried across a split.** Where one incident is split across two dates, the arrest count stays on its own date and the fatalities on theirs. Each row records what happened on its own date.

displaced_count

What it counts. People displaced, recorded on the row for the act that displaced them. Where the note names that act — shelling, an airstrike, a raid — the figure goes on that row, and no second Forced displacement row is opened for it. The field exists on every row whatever the event type, so no figure is lost either way.

WHAT IT DOES NOT COUNT.

- **Households.** A household figure is never converted to people. Where a note gives only households, displaced_count is 0 and displaced_households: N is recorded in conflict_flag.
- **A figure another row already carries.** Where a clash row already records the displacement, no standalone Forced displacement row repeats it.
- **The same flight reported again.** One flight is often reported for weeks, by several outlets and by the same outlet at intervals. Every such row stands, but a row that restates the same

people — the same villages, the same fighting, no new movement — records 0, and the coding record names the row that holds the figure. A row whose note says people moved again, or that villages not named before now fled, records a second movement and carries its own figure.

- **A figure for a wider area.** A figure covering more than the row's township — a highway corridor, a district — is not attributed to the township.
- **Distinct persons.** `displaced_count` counts displacement events, not people. A total is the sum of recorded displacement figures, and one person displaced three times is counted three times. It is not a count of distinct displaced persons and is never presented as one.

defected_count

What it counts. Armed personnel who left one side and joined another. The note must name the side received into — surrendered to, joined, defected to. A soldier or police officer who joins the Civil Disobedience Movement is counted: the opposition calls on them to join it and supports those who do. Only armed personnel are counted. Prison guards count as security forces, and pro-military militia as armed personnel. Family members and civil servants who surrender alongside them are removed where the source gives a breakdown; where it gives a single figure with no breakdown, the figure is recorded as given. A surrender described with no number opens its Defection / Surrender row with `defected_count` 3 and Defection or surrender occurred; number unknown., like any quantity stated without a count (section 3).

WHAT IT DOES NOT COUNT.

- **Desertion, abandonment or repatriation.** A soldier deserting, a garrison abandoning a post, or troops fleeing across a border and being repatriated to their own army have left one side without joining another: the row is Strategic Developments / Forces deployment / Positioning with `defected_count` 0.
- **A defection through a facilitator.** A note saying someone defected with the help of a named group names who helped him leave, not who took him in. Money, a job or safe passage arranged afterwards is support, not enrolment.
- **An administrator resigning.** Under pressure from a named actor this is Coercion / Intimidation; with no actor named it is Other / Other. Both take `defected_count` 0.

released_count

What it counts. People an actor released from its own custody. It fills on the Prisoner release sub-event, and on one other row: where the same note records people detained and released on the same date, as part of the same event, the release figure goes on that row alongside arrested. A release on a later date opens its own Prisoner release row on the date the note gives — with the note's count, or with 3 where it gives none (section 5, rule 4) — and that sub-event is widened beyond amnesty and exchange to hold any release of people an actor was holding.

WHAT IT DOES NOT COUNT.

- **A rescue or extraction.** Where a third party takes them out, or another force recovers them, the holding actor released nobody: the row is Other / Other with `released_count` 0. Counting

it as a release would put a rescue and an amnesty in one field and make the figure unusable as a measure of what detaining actors do.

- **A transfer of custody to another state.** `Other / Other, released_count 0`.
- **A release before the row's own event**, or of people the row does not record as detained.
- **A release the source does not date.** Where the note gives no date and no duration for a release, no `Prisoner release` row opens, whether or not it gives a count, and the arrest row keeps `released_count 0`: the release is not placed on a date the source does not give. The coding record says so — `Release occurred; date not given.`, with the count beside it where the count is known. A release date is never inferred.

A figure split across rows

Where a stated total is split across several rows — across the days of a campaign, or across the locations one strike hit, and only where the note itself states the split — the buckets are split with it where the source attributes them. Where it does not, all three buckets are set to 0 on every row of the split and the breakdown is recorded once, on one row. The full unsplit breakdown is never repeated on more than one row.

A figure repeated across rows

Where a multi-day operation's note carries a recap sentence, that sentence reaches every day's row, and a figure can be extracted onto each of them. The people were taken once. The figure is recorded on the row where the note first reports it and set to 0 on every other row of the operation, with `conflict_flag` on each naming where it is held. A cumulative total restated on a later day's row is never additional, and a combined figure — 58 detained plus 21 surrendered recorded as 79 — is the same two quantities again, not a third.

Casualties in a campaign go on the row whose date the note attaches them to. Where the note gives only a total for the span, it goes on the last row of the span and every other row carries 0; it is never spread across the rows or repeated on them.

Zeroes and multi-event source notes

A figure absent from a single row does not mean it is absent from the dataset.

A single source report often describes multiple distinct events — an arrest and an arson, two protests, or artillery shelling across five townships. MCARI records each event as a separate row but applies the same source note to all of them. Consequently, if a figure is mentioned once in that note, it is recorded in only one of those rows and is entered as zero in the sibling rows.

A zero in a numeric field does not indicate a failure to capture data; it means that row is not the one holding the figure. Before concluding that a figure is missing, every row sharing the same source note is read and the field checked across all of them.

In a verification of 114 rows showing a zero despite their note containing a figure, 64 already held the figure — either on the row itself or on a sibling. Filling those zeroes would have double- or triple-counted the same individuals.

This is the inverse of the rule that a figure is recorded once and set to zero elsewhere. The first prevents double-counting; this prevents restoring a figure that was never missing.

Arithmetic

Figures are never averaged, divided or estimated, and never added — except where the note itself records the addition as its own coding decision. Subtraction has one exception: a stated exact total minus a stated exact component leaves a determined remainder. An exact total is required. "About 40 arrested, all released except 6" gives 34. "More than 60 abducted, one killed and two tortured" gives nothing.

Discrepant figures between sources

Where sources report different figures for the same event, MCARI does not average them, split the difference, or re-derive a figure from the sources. Instead:

1. Where the event note names its sources and records which figure was taken and why, that decision governs. For example: "Myanmar Now reported 6, Irrawaddy at least 3; fatalities coded as 3 for the lowest estimate."
2. A figure resting on a single conflict party's own claim, however specific, does not displace that decision.
3. Where the note records no such decision, or treats the figures equally, the lowest figure any source gives is taken — the one figure every source supports. Myanmar Now 6, Irrawaddy at least 3 and Mizzima 2 gives 2.
4. The figure not taken is recorded in `conflict_flag` with its source named, so the disagreement is preserved rather than resolved away.

The lowest figure settles how many, not whether. It is for sources that agree the event happened and differ on the number.

- A source that gives no figure is not reporting zero, and takes no part in the comparison.
- A conflict party's denial does not overturn independent outlets. The lowest figure among the sources reporting the event is taken, and the denial is recorded in `conflict_flag`.
- Where an independent outlet states that nobody was killed, the case is not resolved in coding.

One outlet revising its own figure. Where an outlet reports a figure for an event and later revises it — four dead in the morning, five by the evening — its latest figure is that outlet's figure, and it enters the comparison between outlets above once. Three limits bound this:

- **Never added.** The latest figure is taken once, never the sum of the versions.
- **One event only.** A later report that takes in further fighting on later days is not a revision: it is a new event, or a total over a period, which is not recorded (section 3).
- **One outlet only.** Figures from different outlets are competing figures, decided by the rules above, never by which came last.

One outlet's running total for a single event — 600, 700, 1,000, 1,198 and 1,233, reported across seven rows — gave more than 6,000 when the rows were summed, against a latest figure of 1,233.

4 The three fatality buckets

Every death recorded in `fatalities` is also placed in one of three buckets, or in none of them. The buckets are `fatalities_military_soldiers`, `fatalities_non_military_armed_groups` and `fatalities_civilians`. A death the source does not attribute stays outside all three, which is the fourth state described at the end of this section.

The two tests

Two questions decide the bucket, in this order:

1. **Which side was the person on?**
2. **Did they belong to an armed formation?**

Whether they wore a uniform decides nothing. An administrator carrying a weapon for personal security is still an administrator: the note must place the person in an armed formation for the military bucket to apply. Detainees who seize weapons are not an armed formation either — prisoners who take a guard's weapon in an escape, or people who pick up arms in the course of being held, remain outside an armed formation for this purpose. The bucket asks whether the person was a member of an armed formation, not whether a weapon was in their hands when they died.

`fatalities_military_soldiers`

What it counts. Deaths of the coup side's armed forces: the Myanmar Military, the Police, the BGF, Pyu Saw Htee and other pro-military militias.

What it does not count. A person aligned with that side who was not in an armed formation. An administrator, an informant, a USDP member or an official appointed after the coup is a civilian for this field, armed or not. It also does not hold a death whose side the note does not give.

`fatalities_non_military_armed_groups`

What it counts. Deaths of the armed opposition: PDF units, EAOs and local defence forces.

What it does not count. Unarmed officials, organisers and supporters on the opposition side, who are civilians for this field. Nor a death whose side the note leaves unstated.

`fatalities_civilians`

What it counts. Everyone else. This includes administrators, informants, party members and other people killed for a role, where they were not members of an armed formation.

What it does not count. A member of an armed formation on either side. Nor a death the source declines to attribute.

A convention figure takes a bucket like any other

Where a note states that people were killed but gives no number, `fatalities` takes the convention figure of 3. That 3 is placed by the same two tests as any other death. Where the note says which side the dead were on — "some Myanmar Military soldiers were killed" — the 3 goes to that side's bucket. Where it does not, all three buckets stay 0. Where the note names more than one side but gives no count, the three buckets stay 0 and the figure sits in `fatalities` alone — the note does not say how the 3 divides, and MCARI does not divide it.

The actor string does not decide the bucket

The bucket follows the side the note gives, not the name in the actor field. PMF and People's Militia Force are used in this dataset for both pro-military and pro-opposition militias, so the string alone cannot settle it: read which side the note places the group on. Pro-military deaths go to `fatalities_military_soldiers`, deaths of a group fighting the military to `fatalities_non_military_armed_groups`. Where the note does not say, all three buckets stay at 0 and the note's wording is kept in `conflict_flag`. A stated affiliation is trusted.

Why a person killed for a role is a civilian

An administrator or an informant killed for their role is recorded in `fatalities_civilians`, not with the armed opposition. The alternative — treating them as aligned with a party to the conflict rather than uninvolved — fails for two reasons. It would leave one bucket holding pro-military militias, opposition fighters and unarmed officials at once, so nothing could be asked of it. And the fact it would preserve — that the person was killed for a role — is already carried by the event type `Targeted Violence / Targeted killing`, which is where it belongs.

Not attributed — the fourth state

The three buckets do not have to sum to `fatalities`. Where a source gives a combined figure and states no split, the buckets stay at 0 and the source's own wording is kept in `conflict_flag`. Where a stated total is split across several rows and the source does not attribute it per row, every row of the split carries 0 in all three buckets, and the breakdown is recorded once, on one row, in a form that names the total and where it was split. The unsplit breakdown is never repeated on more than one row: repeating it would make the buckets sum to more deaths than occurred, while zeroing them costs only detail the source never supplied. Either way the `fatalities` total stays exact.

What the buckets never do is exceed their own row. The sum of the three describes the deaths recorded on that row, and a bucket larger than the row's own `fatalities` is an error whatever produced it.

5 Five coding rules that turn a note into a number

1 • The person-versus-place test

This is the distinction that decides most disputed rows.

What was attacked	Result
A specific person — shot in the street, killed at home, ambushed on the road	Targeted killing
A place or position — police station, checkpoint, booth, outpost, base, even attacked one-sidedly	Armed clash, or by method — Landmine / IED, Grenade attack

The test is counterfactual: if that qualifying-role person had not been there, would this attack still have happened? Yes, and the target was a place, position, unit or convoy — the event is `Armed Conflict`, coded by method, and the person's death is incidental. No, and the attack existed because of that person — the event is `Targeted Violence`.

A police post attacked one-sidedly is not a targeted killing, even though the Police are a qualifying role, because the post would have been attacked whether or not any particular officer was inside.

Vehicles are the case that recurs. A military or police vehicle, or any vehicle whose occupants are security personnel travelling as such, is the mobile equivalent of a position: attacking it is `Armed Conflict`, coded by method, and `Armed clash` where no method is stated. A private or civilian vehicle attacked because a qualifying-role person is inside it is person-targeting, and the row is `Targeted killing` — a private car is not a target until someone knows who is in it. Where the note does not say what kind of vehicle it was, the default is `Armed Conflict`. A mine or IED initiating the attack overrides all of this and stays `Landmine / IED` whoever was inside.

2 • Minimum-confirmed quantities

A vague quantity is a quantity, not a missing number. The wordings and the figure each one takes are set out in section 3 under "Quantifying vague terms". A note saying "thousands fled" does not take 0 and a standing wording; it takes the figure the table assigns. Writing 0 there would say the source gave no quantity, which is false, and the figure would be lost.

A lower bound is a minimum and the minimum is the value: "more than 30 arrested" fills `arrested` with 30, not 0 and not a blank. What a lower bound cannot do is anchor a subtraction. "More than 30 arrested, four of them released" gives `arrested` 30 and `released_count` 4, and never "26 still held" — subtracting from an open bound gives another open bound, not a figure.

Where a field holds a floor, the standing wording for an uncounted quantity is not written on it. `Arrests occurred; number unknown.` on a row carrying `arrested` 30 contradicts the row. The floor is recorded in `conflict_flag` instead — "more than 30 reported; recorded at the floor of 30".

3 · A figure already coded stays coded

This is not the table above doing its work twice. The table derives a figure from a word; this rule protects a figure already derived from being withdrawn to 0 by a later pass.

Where a row carries a "coded as N" figure — in the note or in `conflict_flag` — that figure is a recorded coding decision and it stays. A row reading "Total arrested coded as 13 (3 in the morning + at least 10 at night)" records 13; a row reading "Unknown fatalities coded as 3" records 3. A figure is withdrawn as fabricated only where `conflict_flag` shows it was never sourced and no convention supplies it.

Misreading this rule once caused 108 rows to be withdrawn to 0 and restored from backup.

It is not an instruction to add a figure to a row that lacks one.

4 · "It happened but was not counted"

Where a note establishes that a quantity occurred but gives no number, the field takes 3 — see the table in section 3 — and `conflict_flag` carries the standing wording for that quantity, kept word for word so that a later sweep finds every row holding one:

```
Unknown fatalities coded as 3. · Arrests occurred; number unknown. · Displacement
occurred; number unknown. · Defection or surrender occurred; number unknown. · Release
occurred; number unknown.
```

A sixth wording stands apart from those five. `Casualties unknown.` records that the source never said whether anyone died, and the field it settles stays at 0.

Other forms appear in reporting and mean the same thing — "the number is not reported", a bare "not reported". The canonical wording above is what goes in `conflict_flag`, not the note's variant. The wordings are fixed so that one sweep finds them all; a variant of a coder's own making is invisible to that sweep.

Two limits bound the rule. Each wording resolves only its own quantity: a wording settling fatalities says nothing about arrested or displaced count. And where the note says plainly that nobody was harmed — "no one was hurt", "the villagers had already moved away" — the field is 0 and `Casualties unknown.` is not written either. Nothing is unknown; the note answered it.

5 · A clash and its civilian casualty kept on one row

An armed clash and the civilians killed in it are one event, recorded on one row, with the deaths in `fatalities_civilians`. Deliberate burning or killing after a clash is a separate event and stays separate. A mine detonation followed by return fire is one row, `Landmine / IED`; a mine or IED with a civilian victim and no combat is not split either.

Three further rules follow the same precedence — the row that holds the death is the row the person was killed on, and splitting would count one person twice:

- Sexual violence where the same victim was also killed is one row, `Sexual violence`, with the fatality recorded there.

- Where a person used as a human shield is killed in the course of the event being coded, the fatality goes on the Human Shields row and no separate Attack on civilians row opens.
- Deaths in a fire an actor set are one act: the row is Attack on civilians with the fatality on it, not a separate Arson / Property destruction row.

The date is what bounds all of these. Where both things happened on the same date, one row covers them — the people detained and the people harmed are the same people. Where the dates differ, the event is split, and neither figure is carried across: each row records what happened on its own date.

Where the sibling-row rule is published

A sixth rule sits with these in practice: a figure absent from a row is not a figure absent from the dataset. It is published in section 3, under "Zeroes and multi-event source notes", because it governs where a figure sits rather than how an event is classified.

6 Worked examples

The examples here are chosen because they are borderline: each is a case where two categories could be argued, and the rule that settles it is the point of the example. Each gives the event as the note records it, the question a coder faces, the rule that decides it, and the row that results.

Example 1 · MCARI-03292 — a meeting attacked, and no victim named

The note. On 19 March 2021, in Za Yat Hpyu Kone village, Myawaddy Township, Kayin State, an unidentified armed group attacked village administrators and people's militia members holding a meeting. Two administrators died at the scene; a militia member later died in hospital.

The question. Three people were killed and none is named. Read as an attack on a gathering, the row is *Violence against Civilians / Attack on civilians*. Read as an attack on the people at it, it is *Targeted Violence / Targeted killing*.

The rule. A name is not what makes a killing targeted; a qualifying role is. Village administrators at any level, and members of a pro-military militia, both hold qualifying roles (section 2). The test is counterfactual (section 5): if those people had not held those roles, would the meeting have been attacked? The note gives no other reason for it.

THE ROW.

```
2021-03-19 · Targeted Violence / Targeted killing
actor1 Unidentified Armed Group · assoc_actor_1 -
actor2 Administrator · assoc_actor_2 PMF
fatalities 3 · civilians 2 · military and pro-coup 1
```

Both groups at the meeting were attacked, so both are on the target side: Administrator leads and PMF sits with it in *assoc_actor_2*.

The militia member is counted in *fatalities_military_soldiers*, not with the two administrators. The buckets are decided by whether the person belonged to an armed formation, not by which side the attack was aimed at (section 4).

Example 2 · MCARI-04280 — five police detained during a clash

The note. On 5 April 2021, in Pinlebu town, Sagaing Region, police and around forty soldiers fired into a ward at about 10 pm. Residents blocked the road and returned fire with hunting rifles. The exchange ran through the night. Two civilians were killed, and five police officers were detained by residents.

The question. The detentions could be recorded on the clash row in *arrested*, or opened as a separate *Coercion / Arrests / Detention* event.

The rule. Capture that happens while fighting is under way is an outcome of the clash, not an act of its own: **the line is timing, not status** (section 2). Detention becomes its own row only where the note places the taking after the person has broken away from the engagement.

THE ROW.

```
2021-04-05 · Armed Conflict / Armed clash
actor1 Myanmar Military · assoc_actor_1 Police
actor2 Unidentified Armed Group · assoc_actor_2 -
fatalities 2 · civilians 2 · arrested 5
```

Police and soldiers fired together, so both are on the acting side: the military leads and Police sits with it in `assoc_actor_1`.

That the captives were police and the captors residents does not move the row. The rule turns on when they were taken, not on who held whom.

Example 3 · MCARI-08789 / 08796 — a death in custody, a day after the arrest

The note. In Bae In ward, Kale town, Sagaing Region, military forces arrested a 15-year-old boy on 19 June 2021. He died in custody on 20 June. One other person was arrested on 20 June.

The question. One row, or two.

The rule. A detainee who dies in custody never opens a second row: the death is recorded on the row that already exists for the detention. That holds while the arrest and the death fall on the same date. Where they fall on different dates, they are two events, each recorded on its own date (section 2).

THE ROW.

```
2021-06-19 · Coercion / Arrests / Detention arrested 1
2021-06-20 · Violence against Civilians / Attack on civilians fatalities 1 · civilians 1 · arrested 1
```

The `arrested 1` on 20 June is the other person taken that day. It is not the boy counted a second time.

7 Geography

The three administrative levels

Every row is placed as precisely as its source allows, across three fields:

Field	Level
admin1	state or region
admin2	district
admin3	township

admin3 is the working level of the dataset and the one most analysis groups by. The naming follows Myanmar's administrative hierarchy and MIMU's P-code scheme, in which township is admin3 and district is admin2.

The reference table

Geography is resolved against the **MIMU P-code table, version 9.6**: 358 townships, each with its state or region, its district and its P-code. Place names below township level are resolved through the MIMU gazetteer, 72,569 villages and 3,524 wards.

admin1 is spelled as the dataset spells it rather than in MIMU's form — Bago-East, Bago-West, Shan-North, Shan-South, Shan-East, Naypyitaw. A cross-border event carries the country in admin1, with the remaining geography fields blank.

The two-key join

admin2 and pcode_admin3 are derived from admin3 through the reference table, joined on admin1 **and** admin3 together. The second key is not redundant: Htantabin and Minhla each name two townships in two different regions, and a join on the name alone picks one of them silently.

admin3 — the order of preference

admin3 is taken from the note, in a fixed order:

1. The township the note names.
2. Failing that, a village, ward, market, road, school, monastery or other identifiable place the note names, resolved to its township through the gazetteer. **The resolution is accepted only where it returns exactly one township within that admin1.** A name returning two or more is not resolved and falls through to the next step: 11.9% of place names are ambiguous within their own region, and Aung Chan Thar alone names villages in 21 Sagaing townships.
3. Failing that, the town or city the note identifies.
4. Failing that, a nationwide or multi-region event takes the seat it was announced from — in practice Naypyitaw — or no township, with geo_precision 4. coded as ... is a note convention saying where a multi-place event was filed, and is never an admin3 value.

A township name the reference table does not carry — a district name one armed group uses, a town that administers its surroundings in practice — counts as no township named, and step 2 decides it: the village the note names, under the same exactly-one condition within `admin1`. The coding record names the place that produced the township. Once a town has been resolved this way, a later row naming the same town takes the same township, and its coding record names the rows the mapping came from. Such mappings are held in a decided file of their own, the town-to-township mapping file (`mcari_town_township_mappings.csv`), beside the reference table; a town enters it only by MCARI's decision, never because a coder found a likely match.

The place name always comes from the note; only the township it sits in comes from the gazetteer. `admin3` is never inferred from anything outside the note, except a town-to-township mapping already established from another row's note.

`admin2` is derived, never read

`admin2` holds the district and is never taken from the note, even where the note names one. It is produced by the join above. Where `admin3` holds something other than a township — a city or a district — the derivation has nothing to join on and `admin2` is left blank. **Blank is that row's correct final value, not a gap.** A district is never guessed from a city name. `admin3` keeps whatever the preference order gave it, and `geo_precision` records the level reached.

`PCODE_ADMIN3` is the MIMU P-code for the township in `admin3`, and comes from the same reference row as `admin2`. It is blank wherever `admin2` is blank.

latitude and longitude

latitude and longitude give the place the source names as the event location, not the township's centre. The point is taken from the MIMU P-code workbook v9.6 — the village sheet for villages, the town sheet for towns — by reading the note, extracting the place it names, and looking that place up within the row's own township. A coordinate is recorded only where the lookup returns exactly one candidate. Coordinates are recorded at the precision the workbook carried, and that precision varies from entry to entry.

Where the note names no place inside the township, where the name matches more than one place within it, or where the note's location clause contradicts itself, the fields are left blank. A township or district centroid is never substituted. A blank records that the event's position within the township is unknown; a centroid would misrepresent it as known.

A small number of rows carry a coordinate that does not sit inside its own administrative fields. These are Forced displacement events, which MCARI codes to the township people were displaced from. Where the source names only the place they sheltered in, the coordinate marks that place.

`geo_precision` — the level actually reached

`geo_precision` records the administrative level `admin3` reached, so that grouping by `admin3` does not silently mix a township of 40,000 people with a city of 1.2 million. Every row carries a value.

Value Meaning

1	township — admin2 derived, pcode_admin3 present; or a township the MIMU reference does not carry and no place in the note resolves, held in admin3 as named with admin2 and pcode_admin3 blank. A town the note calls a town is 2.
2	city or town, no township resolved
3	district or region only
4	nationwide or multi-region, coded to the seat it was announced from — in practice Naypyitaw — or left without a township

Most rows are placed to township level. A blank admin2 does not force the value down: the value is the level the row is located to, not whether a particular column is filled. A cross-border row is 3: with only admin1 filled, the level reached is the region, however precisely the note places the event.

An event in more than one township

Where a note names two or more townships, the first question is whether it describes one event or two. Two protests, one in Hlaing Tharyar and one in Lanmadaw, are two events and take two rows, each with its own admin3. A single event that genuinely spans townships — a clash on a boundary, a march passing through several — stays one row, coded to the **primary township**, the one where it began or where its focus was, with the others named in the note. A multi-township event is never pooled up to the district or city level: that would collapse distinct events and destroy the detail the discrete-event model exists to keep.

A township is not required for a row to exist

Where the named place cannot be resolved to a township — a district, a city, a river, a stretch of coast — the row is still recorded, with the place it does have and the precision level that describes it. Dropping such rows would lose real incidents that no other source reported.

A blank is for a source that does not give a township, never for one that is merely hard to find. The reference table and its alias list are searched first, on the Myanmar name as well as the English, and the blank is what remains after that. What is still required is a place: an event the source locates nowhere at all is not recorded.

8 Actors

Four of the file's columns name the parties to an event. A group's name in those columns is not the wording a source happened to use: it is a standardised value, and the same group carries the same string in every row of the series. This section sets out what the four fields hold, how a name is decided, and what the two classifying attributes — actor type and alignment — mean.

The four fields

Field	What it holds
actor1	who carried the act out
actor2	the primary target or opposing force
assoc_actor_1	others on actor1's side
assoc_actor_2	others on actor2's side

actor2 is blank where the event has no target or opposing group — sabotage, an attack on inert property, a one-sided declaration, a troop movement. The blank turns on the absence of a target, not on the sub-event's name, and it is a correct final value rather than a missing one.

Two sub-events fix their own actors. On a *Peaceful protest* row actor1 is *Protesters*, or *Pro-Military Protesters* where the note establishes a pro-military turnout. On a *Protest with intervention* row the intervening force leads, because the row records the act of breaking up the protest rather than the protest itself: actor1 is the force that intervened and actor2 is *Protesters*. Where a note names both police and military, *Police* takes actor1 and *Myanmar Military* goes in *assoc_actor_1* — the one sub-event in the taxonomy where the police lead over the army. Where a note names only one force, that force is actor1 and no assoc value is invented. Victims of a dispersal stay *Protesters* in actor2 rather than becoming *Civilians*. This orientation does not reach *Peaceful protest*, which keeps *Protesters* in actor1, and it reaches no other sub-event: the sub-event decides, never the note's mention of a force — a row is not recoded to put the police first because police appear in its note.

Three other sub-events fix a field in the same way. On a *Mob violence* row actor1 is *Rioters*, the value for a civilian crowd rather than an organised armed group. On a *Communications blackout* row actor1 is the authority imposing the blackout and actor2 is *Civilians*. On a *Political alliances* row the parties to the alliance each keep their own field: actor2 is blank and the remaining groups are listed in *assoc_actor_1*.

Anything standing with the protesters goes in the *assoc* field on the protesters' own side. On *Peaceful protest*, where *Protesters* is actor1, that is *assoc_actor_1*. On *Protest with intervention*, where *Protesters* is actor2, it is *assoc_actor_2*. Journalists, bystanders and others struck at the protest who were not taking part take *Civilians* in the same field, and *CDM* goes there too where the note places it at the protest.

Named strike committees, students' unions, associations, and occupational or identity groups are never written into an *assoc* field. They stay in the note. The *assoc* fields hold only groups the note places in the act, which is what keeps them usable as the basis of network and alliance analysis.

Where two or more groups act together and the joint force has its own name, that name is used. Where it does not, the actor field carries `Combined Group` and every constituent group is listed in the corresponding associate field under its own standard name. No lead group is chosen. That value is only ever a joint force acting together in a single act: it is not used for state forces acting together, for two classes of victim, for one person holding two roles, or for parties that met or faced each other rather than fought alongside each other.

How a name is standardised

Every written form a source may use is held as an **alias**, and each alias resolves to one **standard name**, which is the value written into the file. The source's own wording is never discarded. This is a data-integrity rule rather than a formatting preference: it preserves the trail from what a source wrote to what MCARI recorded, and it allows a standard name to be revised later without losing the ability to match the older wording.

The naming rules that matter to anyone reading the field:

- **Exactly four abbreviations are used in a standard name** — PDF, PDT, LDF, CDF. Every other force term is part of a group's own name and is never abbreviated or altered. `Kamma Defense Force`, `Yaw Defense Force` and `Cobra Column` stay as written.
- **Place names are never abbreviated.** LDF (Myinmu) is correct and LDF (MMU) is not, because two different places can share an abbreviation and the group then becomes unidentifiable.
- **A name that identifies one group on its own stands alone.** `Cobra Column`, `Zero Guerrilla Force` and `Dark Shadow` take no type prefix and no location suffix. Only a generic name that cannot identify a group by itself is disambiguated, with the locality in a trailing parenthetical.
- **The locality is taken at the level the name itself states.** A district form is never converted to a township form or the reverse, and a locality is never taken from the row's own geography: a group raised in one region can fight in another, so the event's place is not the group's.
- **An ethnic armed organisation's political and military wings are held as one compound string**, joined by a slash — KIO/KIA, KNU/KNLA, ULA/AA, PSLF/TNLA. The list of these compounds is closed. A slash appears only inside a single organisation's own name; two distinct groups acting together are separated by a semicolon.
- **Unit, battalion and column numbers never enter an actor field.** Numerals that are part of a group's own name stay: `27 Revolution Forces (Myingyan)`, `Area 71 Local Defense Force (Monywa)`.
- **A name is never translated in either direction.** Where a source writes a group's name in English the English stands; where it writes it in Burmese romanisation the Burmese stands.

actor_type

Actor type is a closed list of sixteen values. It classifies a formation's structure, not its side and not its name.

Value	What it covers
State forces	the Myanmar Military, Police, Border Guard Force, Border Guard Police, Prison Guards and Military Security Affairs
Pro-military militia	militias raised and armed in support of the military, including Pyu Saw Htee and the people's militia forces
Ethnic armed organization	an ethnic armed organisation established before February 2021
EA0 splinter	a formation that broke away from an ethnic armed organisation
PDF	a battalion-structured force under the NUG Ministry of Defence
NUG structure team	a civil or security body of the National Unity Government, at any level
Ethnic defence force	an ethnic-based defence force formed after February 2021
Local defence force	a local armed group under neither the NUG nor an ethnic armed organisation
Urban guerrilla	a clandestine armed group operating in an urban area
Specialist unit	a unit organised around a capability rather than a territory — drone, demolition, technical
Non-EA0 armed group	an armed group that is neither an ethnic armed organisation nor a PDF
Political body	a political or administrative body
Foreign state force	the armed forces of another state
Joint operation	a formation made up of two or more groups acting together
Civilian	civilians, protesters and rioters
Unidentified	a group whose name and type are both unestablished

A people's militia is written PMF in the four actor fields, whatever it calls itself, and its own name is held in the registry alone. Pyu Saw Htee and other named pro-military formations keep their own names. The type is the same for all of them.

Type, side and name are three separate things, and the separation is deliberate. A group can change sides without changing structure, and a group's own name may carry a force term that does not describe what it is — a formation calling itself a PDF may be a local defence force. Putting the type inside the name would mean renaming a group whenever it reorganised, which breaks the series; holding it in its own attribute means one value changes and the name stays.

Nothing is folded into a name that an attribute can carry. A successor body, a renamed formation or a group that changes sides is held under its own standard name, with the relationship carried by type, alignment and the registry's own notes. Folding two names together to make a category easier to count destroys a distinction permanently; keeping them apart and typing them alike costs nothing and keeps both.

A type is left blank where no source establishes what a group is, and where a formation is not a fighting formation at all — a named medical, ambulance or nursing unit travelling with a column takes no type, and no seventeenth value is added for it.

alignment

Alignment is one of `Pro-SAC`, `Anti-SAC` or `Other`, and it records the side a group took rather than what it is.

It follows from the actor type and is then overridden on named exceptions. State forces and Pro-military militia are Pro-SAC. The resistance types — PDF, NUG structure team, Ethnic defence force, Local defence force, Urban guerrilla, Specialist unit, Joint operation, Non-EAO armed group — are Anti-SAC. Civilian and Unidentified carry no alignment. Ethnic armed organisations are set individually, because their positions differ and several of them fought both the military and each other.

A group carries one alignment for the whole series. Where a group changed sides during the period, the value is the side it held for most of the span, and the change is recorded in the registry's notes. Anyone grouping events by alignment across time should know that the attribute is a single value per group and not a per-event judgment.

Where a name cannot be resolved

Where the correct standard name is genuinely unclear, **the raw string from the source stands** and what is unclear is recorded against it. Nothing is inferred and no name is constructed from a plausible reading. MCARI uses groups' real names where they can be established; it does not invent tidy ones. A bare or unresolved form in an actor field therefore means the reporting did not establish which formation acted, not that the value was overlooked.

Where reporting names no group at all, the field is not left empty: an unidentified perpetrator is still an actor. Unidentified Armed Group is used where an armed group acted, Unidentified Actor for a single unnamed person, and Unknown allies for unnamed allies of a named actor. A perpetrator a source only hedges at is Unidentified Armed Group; the hedged name stays in the note and goes no further.

Names that are never folded together

Two organisations that share an abbreviation are never merged, however similar the strings look. The KNLA in KNLP/KNLA is the Kayan New Land Army and not the Karen National Liberation Army, so the bare abbreviation is never registered on its own. ANC/AA is a separate organisation from ULA/AA and is never folded into it. Where one generic name covered two unrelated groups, the name is split and each is disambiguated by its locality, and the split is made only where reporting establishes that the groups are genuinely different — not merely that a name appears in two places.

The same rule governs the pair that is easiest to confuse: a NUG structure team and a Ministry of Defence battalion formation in the same township may both be written PDF (place) by reporting. They are held apart as PDT (place) and PDF (place), because the name itself has to separate them once both reach the file.

9 Limitations

MCARI records what reporting establishes. What follows is what the data therefore cannot show. It is set out here rather than left for a reader to discover, because a figure whose limits are unstated will be used beyond them.

The record is built from reporting

Where reporting is thin, the record is thin. Areas closed to journalists, periods of communications blackout, and places far from any outlet's reach produce fewer recorded events than places under continuous coverage — not because less happened there, but because less of it was reported.

A zero in MCARI's data means nothing was recorded. It does not mean nothing happened. This applies to a township with no events in a quarter as much as to a numeric field reading 0 on a single row.

Coverage is also uneven across the period. The volume of Myanmar reporting, the number of outlets operating, and their access have all changed since February 2021, and a change in the record over time may reflect a change in reporting as much as a change in the conflict.

The counts are minimums

Where a source reports that people were killed, detained or displaced without giving a number, MCARI records a fixed conventional figure rather than an estimate. Where a source gives a quantity in words — hundreds, thousands — the value recorded is the least those words allow. Both are floors. **The real figure is higher than the recorded one more often than it is lower, and by an unknown margin.**

Where a death is reported without saying whose it was, it is counted in the fatality total and in none of the three side buckets. The buckets therefore sum to less than the total, and the difference is stated wherever the buckets are published.

Nothing is estimated, interpolated or modelled. A quantity MCARI cannot establish is not supplied by calculation.

What the fields do not count

There is no field for the injured. An injury figure a source gives is recorded in the note, which is not supplied, and appears in no column.

Displacement counts **recorded displacement events, not distinct people**. A person displaced three times appears three times. The figure cannot be read as a population.

Location

A coordinate is recorded only where the place a source names resolves to exactly one candidate inside the row's township. Where it does not, the coordinate fields are blank, and a township or

district centroid is never substituted in their place.

Not every event resolves to a township. Where a source gives only a district, a region, or a nationwide measure, the row still exists and `geo_precision` records the level actually reached.

Township-level analysis must exclude rows below that level, and the count of excluded rows is published with each issue.

Actors

A perpetrator that reporting does not identify is recorded as unidentified. This is common, and it is a property of the reporting rather than a gap in the coding.

Names are standardised against MCARI's registry, so the string in the data is often not the wording the source used. A group's structural type is left blank where no source establishes what the group is.

The source field

The outlets that reported an event are named, but they are not classified by type. **The share of events resting on a conflict party's own account therefore cannot be separated out from the data as supplied.** MCARI checks such claims before recording them; a user of the data cannot repeat that check from the columns alone.

The series is revised

MCARI's dataset is a living record. A row is corrected when better information reaches it, and corrections are not held back to preserve a published figure.

Reports are therefore cited against a frozen edition. A figure published in an issue remains reproducible from the edition that issue names, even after the live dataset has moved past it.

Appendix A The taxonomy at a glance

Seven event types, thirty-six sub-event strings, thirty-seven valid pairs. Shelling / Artillery / Missile attack sits under two types, marked ‡. **A validator checks the pair, not the string.** Definitions are in section 2; the lines here are reminders, not the rule.

Armed Conflict — direct military engagements between armed actors

Armed clash	two-sided fighting, or an attack whose method the source does not name
Airstrike	a strike from a fixed-wing aircraft or helicopter
Drone attack	a strike from an unmanned aircraft
Shelling / Artillery / Missile attack‡	indirect fire from ground-based systems
Landmine / IED	a device emplaced to attack, whoever the victim
Grenade attack	a thrown or directly fired explosive — people, vehicles or property
Sabotage	inert, unoccupied infrastructure destroyed outside combat — a bridge, a pylon, a tower
Siege / Blockade	a place surrounded and cut off

Violence against Civilians — deliberate violence targeting non-combatants

Attack on civilians	ground force, small arms, close violence
Sexual violence	conflict-related sexual violence
Arson / Property destruction	civilian property burned or destroyed
Shelling / Artillery / Missile attack‡	indirect or air-delivered fire on a civilian target
Mob violence	a civilian crowd attacking civilians

Targeted Violence — deliberate killing of specific individuals

Targeted killing	a person attacked for a qualifying role, on either side
------------------	---

Coercion — coercive control exerted over people

Arrests / Detention	detention as the primary act
Enforced disappearance	custody denied or the person's whereabouts concealed
Forced displacement	an actor drove people out
Forced conscription / Porter duty	people taken for military service or labour
Human Shields	civilians taken and used to shield an armed actor
Communications blackout	telecommunications or networks shut down
Intimidation	pressure applied, with no arrest and no injury

Protests / Civil Unrest — political expression people chose to make

Peaceful protest	an assembly with no intervention
Protest with intervention	security forces disperse it by force

Strategic Developments — politically or militarily significant non-violent events

Ceasefire / Peace talks	negotiation or agreement between parties
Territorial control change – Armed opposition	a place captured and held by opposition forces
Territorial control change – Military	a place retaken and held by the military
Headquarters / Base established	a permanent or semi-permanent position established
Political alliances	formations merge, or place themselves under one command
Political statement / Declaration	an official statement, order or declaration
Defection / Surrender	an armed actor leaves one side for another
Prisoner release	people let go by the actor that held them
Forces deployment / Positioning	movement or positioning for a stated period, no combat
Military clearance operation	a place entered with coercive intent, no combat
New armed group / Faction formation	a formation that did not exist before
Weapons seizure	weapons taken outside combat
Military training / Graduation	training held or completed

Other — events that fit no category

Other	harm with no attacker, harm inside one force, or an outcome that resembles a category but is not it
-------	---

The four boundaries that decide most disputed rows

Target first, method second. What was attacked decides the type. How it was attacked chooses the sub-event once the type is settled.

A place, position or unit attacked is Armed Conflict — a post, a checkpoint, a base, a convoy — even where only one side fired. **A person attacked for a qualifying role is Targeted Violence.**

Indirect or air-delivered fire with no opposing combatant named as the target, striking a civilian target, is VaC / Shelling / Artillery / Missile attack — not Attack on civilians, which is for ground force and close violence. Fire that struck neither an opposing force nor a civilian target stays Armed Conflict.

Capture while fighting is under way stays on the clash row, with arrested filled there. Detention opens its own row only where the note places the taking after the person broke away from the engagement.

Appendix B The standing wordings

Where a source establishes that something happened but gives no number, the field takes a conventional figure and the row records which convention was used. The wordings below are the only ones used for that purpose, and they are written word for word, so that every row carrying one can be found in a single pass.

Wording	What the source did	Field	Value
Unknown fatalities coded as 3.	reported a death without counting it	fatalities	3
Casualties unknown.	did not say whether anyone was killed	fatalities	0
Arrests occurred; number unknown.	reported arrests without counting them	arrested	3
Displacement occurred; number unknown.	reported displacement without counting it	displaced_count	3
Defection or surrender occurred; number unknown.	reported a defection or surrender without counting it	defected_count	3
Release occurred; number unknown.	reported a release without counting it	released_count	3
Release occurred; date not given.	reported a release without a date, counted or not; any count is in the coding record	released_count	0

Three limits bound their use.

Each wording settles its own quantity and no other. A row recording that arrests occurred without a count says nothing about how many were displaced; that field is read on its own.

A wording is never written beside a figure the note gives in full. Arrests occurred; number unknown. beside arrested 30 from "30 arrested" contradicts the row, and makes it a permanent false positive for anything searching for uncounted quantities. Where a field adds a stated number to an uncounted group — an owner (1) and employees with no number (3), recorded 4 — the wording marks the uncounted part and stands beside the figure.

Where a source says plainly that nobody was harmed, the field is 0 and no wording is written. Nothing is unknown; the source answered it.

These wordings are not supplied with the data. They are described here because they explain how a published figure was arrived at.