PLOVER

Political Language Ontology for Verifiable Event Records

Event, Actor and Data Interchange Specification

Open Event Data Alliance

`http://openeventdata.org/`

`http://ploverdata.org/`

DRAFT Version: 0.6b2

February 2017

Acknowledgments

Latest update: Wednesday 8th February, 2017 (UTC)

Chapter 1

Introduction

Event data in political science is a structured way of recording interactions between political actors described in text. For instance, a researcher encountering the sentence

> A town in western Sudan's South Kordofan state has been recaptured by Sudanese government forces from the rebel Sudan People's Liberation Army (SPLA).
> (AFP_ENG_19970408.0772)

might want to represent it as ACTOR = "Sudanese military", EVENT = "capture territory", TARGET = "SPLA", a process that can be applied over hundreds of thousands of reports to create datasets of thousands of such events for answering research questions.

The software to extract and categorize these events exists (see OEDA's Petrarch2 (`https://github.com/openeventdata/petrarch2`), for instance), but all event data systems require an ontology defining what actors and events will be recorded and how they will be defined.

Ontologies face difficult tradeoffs: broader groupings of event types and actors are easier to define and implement, are easier to work with, and provide data at a useful level of aggregation, especially analyzed globally. More specialized and specific ontologies will sometimes be necessary for answering certain research questions and allow more granular and subnational study, but are more difficult to implement and may provide distinctions that are not useful to most researchers. In the example above, should the event be categorized as a general "seize" event? Or a more specific "capture territory"? Should the target be coded as a general "Sudanese rebel" or the very specific "SPLA"?

The best level of detail will depend on the question and resources of the researcher. PLOVER is a new event data ontology that choses to be more general, easier to implement (including across languages), easier to use, and at the level of detail demanded by most existing users of event data. At the same time, PLOVER defines and provides guidance on making "PLOVER-compliant extensions" that will fit into the ecosystem of tools for creating and analyzing PLOVER event data.

1.1 Overview

This manual starts with a history of PLOVER and event data for users coming from CAMEO. It then defines the top-level event types that base PLOVER defines. Along with these, it describes the "mode" and "context" fields that PLOVER uses to provide more information about the event and the circumstances in which it's occuring. Next, the manual defines codes for the actors. It then lists the standard fields in a PLOVER event record before describing "PLOVER-complient" extensions that researchers can use for building custom datasets.

1.2 History of PLOVER

During the twentieth century, two coding frameworks dominated event data research: Charles McClelland's WEIS (McClelland, 1967, 1976) and the Conflict and Peace Data Bank (COPDAB) developed by Edward Azar (Azar and Sloan, 1975; Azar, 1980, 1982). Both were created during the Cold War and assumed a "Westphalian-Clausewitzian" political world in which sovereign states reacted to each other primarily through official diplomacy and military threats. While innovative when first created, these coding systems were less than optimal for dealing with contemporary issues such as ethnic conflict, low-intensity violence, organized criminal activity, and multilateral intervention. Furthermore, McClelland (1983, pg. 177) viewed WEIS as only a "first phase"; he certainly did not anticipate that it would continue to be used, with only minor modifications, for four decades. Event coding ontologies, it seems, tend to be rather sticky.

During the early 2000s, the CAMEO framework—Conflict and Mediation Event Observations— was developed,[1] and originally was intended merely to support an NSF-funded project on the study of inter-state conflict mediation. Instead, it was gradually adopted as a "next generation" coding scheme, notably for the DARPA-funded Integrated Conflict Early Warning System (ICEWS) project (O'Brien, 2010) because it corrected some of the long-recognized ambiguities in WEIS and COPDAB, and was explicitly designed both for automated coding and for the detailed coding of sub-state actors.

As event data came into wider use in the 2010s, several problems with the CAMEO standard— which had never been intended to be a "standard" in the first place—became apparent. These included

- Almost all applications of CAMEO event data aggregated to either the 2-digit "cue category" or the even more general "quad category."[2] No one used all 260 codes.

- Nonetheless, users unfamiliar with the data generating process for automated event coding sometimes assumed every code had been equally well implemented.

- The complexity of CAMEO made it almost impossible to generate a comprehensive set of "gold standard records" and human coders had difficulty agreeing on how to consistently distinguish many of the subcategories: this became particularly apparent as efforts were made to implement CAMEO in Spanish and Arabic.

- Newer coding systems provided information such as geolocation and named-entity extraction beyond the original date-source-target-event format and there was no standard for how to include these in the data.

- The continuing emphasis on coding substate activities demonstrated the need for either new categories or contexts to deal, for example, with criminal activity and events such as natural disaster, elections, and parliamentary behavior.

---

[1]The canonical citation for CAMEO is Schrodt et al. (2009), and the most recent version of the manual is found at `http://eventdata.parusanalytics.com/data.dir/cameo.html`. The original event framework was very much the work of Deborah Gerner and Ömür Yilmaz, with contributions by various coders in the Kansas Event Data System project. The CAMEO manual contains an extended discussion of the issues considered in transitioning from WEIS to CAMEO. Additional details on the development of the automated coding underlying CAMEO can be found in Schrodt (2006) or `http://eventdata.parusanalytics.com/utilities.dir/KEDS.History.0611.pdf`.

[2]Verbal cooperation, material cooperation, verbal conflict, and material conflict, with these usually defined exclusively using the 2-digit categories.

Predictably, because of these and other issues, by 2015 CAMEO was beginning to "fork": both the ICEWS data implementing CAMEO in the BBN/Raytheon `ACCENT` coder (https://dataverse.harvard.edu/dataverse/icews) and the Phoenix data implementing it using the Open Event Data Alliance (OEDA) `PETRARCH-2` coder (http://phoenixdata.org/) differed significantly in some of the CAMEO categories compared to the original implementation in the University of Kansas `TABARI` coder.

To address these concerns, an informal group of academic, government and private sector producers and users of event data met and circulated drafts during the fall of 2016 to develop a new, simplified and more flexible event data specification to replace CAMEO: the end product of those deliberations is the document you are reading. The major changes are the following

- CAMEO's numeric categories demanded a high degree of fluency from users (e.g. remembering that that "09" is "Investigate" and "08" is "Yield"). The numeric categories also implied an ordinal scale that was not always approprate: "assault" is more extreme than "coerce", but "investigate" isn't a more intense than "yield"). PLOVER does away with numeric codes to solve both of these problems and instead uses only a text name for each event type.

- A set of standardized names ("fields") are specified for both the core event data fields and for extended information such as geolocation and extracted texts; most of these fields are optional and where available we use existing specifications, for example the http://geonames.org geographical location field names, ISO-3166 country identifiers and ISO-8601 date and time formats, and *Social Conflict Analysis Database* (SCAD) protest contexts. To accomodate this flexibility, PLOVER defines events in a JSON document structure rather than the CSV formats that most social science datasets are available in. A PLOVER R package will make it easier for researchers to work with the PLOVER dataset and custom datasets using the PLOVER ontology.

- Only the 2-digit event "cue categories" have been retained from CAMEO: our hope is that these are sufficiently broad and distinct that one can achieve a reasonably high level of human inter-coder agreement—hence "verifiable"—on the coding categories, and that these distinctions can be consistently implemented, across all categories, in the automated system. These are defined in greater detail than they were in WEIS and CAMEO.

- The CAMEO 01 and 02 categories dealing with comments have been eliminated.[3]

- The CAMEO 08 "YIELD" category has been split into verbal (CONCEDE) and material (RETREAT) components.

- A new category has been added for criminal behavior.

- Much of the detail 3- and 4-digit details is now delegated to an optional *mode* field (as well as *context*): see Section 2.1 for further discussion of this.

- A "context" field is available, along with standard values, to handle contexts such as disease, natural disaster, elections, parliamentary processes and cyber-security.[4]

---

[3]Ironically, this reverses a decision McClelland belatedly made—and later regretted—in the WEIS specification in the 1960s.

[4]In contrast to the Integrated Data for Events Analysis (IDEA) Bond et al. (2003) system which was in use in the early 2000s, we are not treating disease and disaster as events with targets, instead they are a context within which other activity occurs. This, interestingly, echoes an approach used in COPDAB but not retained by WEIS and CAMEO.

- The complexity of substate actor codes has been limited, and the allowable substate modifiers have been substantially simplified.

- Standard optional fields have been defined for some categories, and the "target" is optional in some categories.

- We have converted all of the examples in the CAMEO manual to an initial set of English-language "gold standard records" for validation purposes—these files can be found at `https://github.com/openeventdata/PLOVER/blob/master/PLOVER_GSR_CAMEO.txt`—and we expect to both expand this corpus and extend it to at least Spanish and Arabic cases. This documentation includes the English-language gold standard records from the CAMEO manual.

Because PLOVER is generally a simplification of CAMEO—the new CRIME category is the one exception—our expectation is that it will be relatively easy to transition the existing CAMEO-based coders into this by simply collapsing the two- and three-digit categories; a similar simple recoding will allow older CAMEO-coded data sets to be converted to their PLOVER equivalents. The standardization of the JSON field names—as well as adoption of JSON as the data interchange format—will allow the development of general-purpose utilities that can work with all formats, in contrast with the current proliferation of incompatible CSV and tab-delimited formats.

1.3   Issues we will not be addressing

In the discussions leading to the development of PLOVER, several additional open issues were raised that we have decided to remain agnostic on:

Temporal markup: This is emerging as a major issue, particular among users who are interested in the long-standing objective of automated chronology generators. While there are some significant efforts on this in the NLP community—`http://www.timeml.org/`—we don't feel we currently have the experience required to make recommendations.

Number of events per sentence: Everyone in the community now has a collection of sentences that could be interpreted as containing anywhere from one to a half-dozen or more events. While this is nominally an ontological issue—"the branch of metaphysics dealing with the nature of being"—the ability to resolve it will be highly dependent on specific parser/coder implementations, so we're leaving it at that level. Same for sentence-level vs paragraph- or article-level coding.

De-duplication: There is no consensus on this beyond noting that the widely-used "one-a-day filtering" is probably not a good idea,[5] and it is a topic where there is currently active research and experimentation, so we're leaving it alone.

As a more general caveat, we would very much encourage anyone who is using a PLOVER-coded data set to look carefully at the software used to implement it, particularly with regard to

- The motivations (and resources available) for producing the data, which will often provide information on how much attention was given to various categories.

---

[5]See `http://eventdata.parusanalytics.com/papers.dir/Schrodt.TAD-NYU.EventData.pdf`

- If the coder is dictionary-based, the event and actor dictionaries if these are available. If the coder is classifer-based, the training cases if these are available.

- Recall/precision tradeoffs and how these play out across different news sources: we are not aware of any existing parser/coder that is optimal for everything from a Reuters story written and edited by people trained at Oxford to a BBC radio transcript from a static-filled French radio report out of Goma, DRC quickly translated into English by a non-native speaker of either language.

OEDA was founded on the principle that there is no "one data set to rule them all": different implementations will have different strengths. PLOVER is at least as much a data-interchange format as a coding ontology, and should simplify the ability of the research to check the robustness of results by looking at multiple independent coding efforts. Do it!

1.4   Why "PLOVER"?

Plovers (*Charadriidae*) are a globally-distributed family of short-billed gregarious wading birds who spend their lives frantically poking through endless stretches of sand and muck trying to find something of interest. It is difficult to imagine a better analogy to the process of coding event data.

Chapter 2

Event Categories

General observation (pas 2016-10-19): I've included separate *mode* and/or *context* lists for CONSULT, MOBILIZE, CRIME, COERCE, PROTEST and ASSAULT but in fact we could probably usefully include them for almost all of the categories. In contrast to CAMEO, these are optional, but they do provide a systematic way of providing more detail. However, there is still too much detail specific to mediation in here.

2.1 Event, Mode, and Context

As noted above, most[1] of the detail found in the 3- and 4-digit categories of CAMEO is now found in the *mode* and *context* fields in PLOVER. More generally, PLOVER takes the general purpose "events" of CAMEO (as well as the earlier WEIS, IDEA and COPDAB ontologies) and splits these into three components: generally "$event - mode - context$" corresponds to "$what - how - why$." We anticipate at least four advantages to this

1. The three "$what - how - why$" components are now distinct, whereas various CAMEO sub-categories inconsistently used the *how* and *why* to distinguish between subcategories.

2. We are probably increasing the ability of classifiers—as distinct from parser/coders—to assign *mode* and *context* compared to their ability to assign subcategories.

3. In initial experiments, it appears that the approach *much* easier for humans to code than the hierarchical structure of CAMEO because a human coder can hold most of the relevant categories working memory (well, that and a few tables easily displayed on a screen)[2]

4. Because the words used in differentiate *mode* and *context* are generally very basic, translations of the coding protocols into languages other than English is likely to be easier than translating the subcategory descriptions found in CAMEO.

While both *mode* and *context* will usually take a single value, in some instances multiple values will be appropriate and this is allowed (and preferable to generating multiple events from a single text). Both fields are optional, and unless an "other" category is explicitly specified—this has been

---

[1] "most" because some of the details coded in CAMEO, particularly those dealing specifically with mediation, have been eliminated.

[2] PLOVER coding may also be easier because everything in the *event - mode - context* coding uses words, not numerical codes, so coders will probably be using the parts of the brain (Broca's area) which are specialized for processing words. No known specialized cognitive facility exists for handling some 250 2-to-4-digit codes.

provided in a couple of the event categories to provide compatibility with other coding systems—if no existing values seem appropriate, the field should be left null, though perhaps with some details provided in the *comment* field, particularly when the record is generated using human coding.

In general, verbal activities only have a *context* since their *mode* is just "verbal." The exceptions are CONSULT where the *mode* indicates how the consultation was done, and THREATEN where the *mode* indicates what action is being threatened.

We anticipate that in general—and consistent with earlier event coding schemes—it will be possible to code *mode* from the same sentence used to code the event. *context*, in contrast, will often need to be coded at the paragraph- or document-level: this differs from earlier automated coding, though probably is similar to human-coded data such as COPDAB and BCOW where *context*-like fields were coded.

2.2 AGREE

Agree to, offer, promise, or otherwise indicate willingness or commitment to cooperate. All cooperative actions reported in future tense are also taken to imply intentions.

2.2.1 Requires target: No

2.2.2 Supplementary fields: None

2.3 CONSULT

All consultations and meetings: this includes visiting and hosting visits, as well as meeting at a neutral location, and consultation by phone or other media. Other useful keywords: "Holding talks" and "discussions", "negotiations, bargaining, or discussions". See the discussion in Section 3.2.2 on the treatment of actors in CONSULT events.

2.3.1 Requires target: No

CONSULT events where there is no clear distinction between whether an actor is hosting or visiting, all participants are coded as source actors.

2.3.2 Supplementary fields:

Table 2.1: CONSULT modes

Name	Content
host	Meeting is hosted by source
visit	Meeting is hosted by target
third-party	Meeting is hosted by a third party
multilateral	Meeting occurs in a multilateral context, typically an alliance or IGO
phone	Consultation occurs via phone or some other remote medium

Adapted from CAMEO.

2.4   SUPPORT

Initiate, resume, improve, or expand diplomatic, non-material cooperation; express support for, commend, approve policy, action, or actor. This event form is a verbal act. Use this code only for political, diplomatic, and non-material support, including recognition of newly independent states, new governments that might have come to power through unconventional means, and initiation of diplomatic ties with an entity for the first time, as are actions which ratify, sign, or finalize an agreement or treaty.

2.4.1   Requires target: No

2.4.2   Potential ambiguities

- Formal pardons and amnesties of arrested persons should be coded as CONCEDE; the actual release or exchange of prisoners should be coded as RETREAT.

- Expressions of regret or remorse for an action or situation should be coded as CONCEDE.

- Promises to sign or ratify agreements and treaties are coded as AGREE

- Military cooperation or defense should be coded as COOPERATE with a `military` *context*.

2.4.3   Supplementary fields: None

2.5   CONCEDE

This covers verbal concessions which have no immediate material consequences, including promised of future concessions, including easing of administrative or legal restrictions on persons and organizations, remove curfews, suspending protests, declarations (but not implementations) of ceasefires and withdrawals from territory.

CONCEDE, like the verbal components CAMEO/WEIS predecessor YIELD, is inherently problem since many concessions deal with promises that certain things will *not* happen, or will happen in the distant future (e.g. many policy changes). So, for example, the lifting of a curfew is, effectively, a promise that people will not be arrested for violating the curfew, which itself is not an event. We're treating such concessions as verbal rather than material even though sometimes they have material consequences, e.g. people coming out in the streets after a curfew is lifted. But only if they believe the government. As noted in Section 1.3, PLOVER isn't really set up to deal with these levels of event dependence.

2.5.1   Requires target: No

2.5.2   Supplementary fields: Use at least a few of the CAMEO cases?

2.6   COOPERATE

Initiate, resume, improve, or expand *mutual* material cooperation or exchange, including

- Initiate, resume, improve, or expand economic exchange or cooperation.

- Military exchanges such as joint military games and maneuvers.

- Cooperation on judicial matters, such as extraditions and war crimes.

- Voluntary exchanges or sharing of intelligence and other significant information .

COOPERATE is distinguished from AID because the activity is generally understood to directly benefit both parties, whereas AID is understood to primarily benefit only the recipient.

2.6.1   Requires target: Yes

2.6.2   Supplementary fields: None? Or rely on the general context codes?

2.7   AID

All provisions of providing material aid whose material benefits primarily accrue to the recipient. Examples include:

- Monetary aid and financial guarantees, grants, gifts and credit.

- Military and police assistance including arms and personnel.

- Humanitarian aid such as emergency assistance.

- Asylum, both to persons in its territories (territorial asylum) and diplomatic asylum on the premises of an embassy.

2.7.1   Requires target: Yes

2.7.2   Supplementary fields: None? Or capture the monetary, manpower, and magnitudes from the context?

2.8   RETREAT

RETREAT covers any events—not just military "retreat" from territory—which have an immediate (not simply promised) material consequences, such as the release of prisoners and hostages, repatriation of refugees, the return of confiscated property, allowing the entry of observers, peacekeepers, or humanitarian workers, disarming, observing a ceasefire or otherwise ending active conflicts, and, of course, a military retreat from, or ceding, territory. RETREAT also covers resignations of government officials.

2.8.1   Requires target: No

2.8.2   Supplementary fields: Almost certainly need some

2.9   INVESTIGATE

All investigations, including those of historical cases. Examples include investigations of criminal activity (theft, killing, etc) and corruption, human rights abuses, war crime, and violations of basic freedoms, military activities such as violations of ceasefire, seizures, and invasions.

2.9.1   Requires target: No

2.9.2   Supplementary fields: None?

2.10 DEMAND

All demands and orders. Demands are stronger or more forceful than a request or appeal—which is not coded in PLOVER—and potentially carry more serious repercussions, although not as much as threats. Coding will need to rely primarily on the language used by reporters to make this distinction. All demands are verbal acts.

Examples from the CAMEO manual include:

- Demand that target engages in some form of material or economic exchange or assistance.

- Demand that target engages in or expands military relations or assistance.

- Demand that target engages in or expands cooperation in judicial matters.

- Demand that target exchanges intelligence or information.

- Demand expansion of diplomatic ties or non-tangible support on particular policies.

- Demands by refugees to be let into the territories of other countries (which should be coded as targets) and asylum demands all fit here. These are not necessarily verbal acts; refugees could be actively seeking shelter or refuge in target countries or regions.

- Demand that the target provides military protection or peacekeeping forces for itself or on behalf of another party.

- Demand for elections, changes in leadership or regime, or constitutional/institutional/policy change.

- Demand provision or expansion of social, political, or other rights, as well as demands for provision of compensation for previously violated rights.

- Require, demand major institutional, constitutional, or regime change.

- Demands for fundamental changes in the political system (e.g. democratization) as well as more limited institutional changes (e.g. changing electoral law).

- Demand that target relaxes administrative restrictions.

- Demand that target stops political protest activities.

- Demand that target releases persons (e.g. prisoners, hostages) or property.

- Demand that target lifts or eases economic sanctions, boycott, or embargo.

- Demand that target allow access to international actors, such as observers, humanitarian agencies, and peacekeeping forces.

- Demand that target stops fighting or takes measures to ease military conflict or tension, for example ceasefires, military withdrawals, and demobilization.

- Order party(ies) to meet, negotiate; this event form can be initiated by either the adversaries or other third parties.

- Order parties to a conflict to reach a settlement, agreement, or resolution of conflict.

- Demand that a third party mediates a conflict or that adversaries accept mediation of another party.

2.10.1 Requires target: No

2.10.2 Potential ambiguities

- This category only applies to verbal demands: demands that take the form of demonstrations, protests, etc. are coded as PROTEST.

- When one or more parties to a conflict call for ending the conflict, that is taken to be an expression of intent on the part of that source actor and is thus coded as AGREE.

2.10.3 Supplementary fields: None? – obviously a lot of *context* fodder in that list...

2.11   DISAPPROVE

Express disapprovals, objections, and complaints; condemn, decry a policy or an action; criticize, defame, denigrate responsible parties.

Examples from the CAMEO manual include:

- Allege, charge the target with, or blame for engaging in crime or corruption.

- Allege, charge the target with, or blame for human rights violations, such as arbitrary detentions for prosecutions, torture, and slavery.

- Allege, charge the target with, or blame for initiating hostilities or engaging in questionable or unjustifiable military actions such as violations of ceasefire, or with war crimes.

- Allege, charge the target with, or blame for spying, espionage, or treason.

- Solicit other parties to take actions against the target.

- Written and institutionalized protests, appeals, and all petition drives and recalls.

- Sue, file civil or criminal lawsuit at domestic or international courts. Source must be the plaintiff or the state, and target must be the defendant.

- Find guilty or liable at a court of law. Source must be the court in question, which could be domestic or international, and target must be the defendant. This event form refers typically to rulings against non-individuals, where imprisonment is not an issue. When individuals are found guilty and are therefore detained, use COERCE instead.

2.11.1   Requires target: No

2.11.2   Supplementary fields: None?

2.12 REJECT

All rejections and refusals. Examples from the CAMEO manual include:

- Refuse to engage in or expand material exchange. Note the difference between refusing to establish or expand material cooperation and reducing or eliminating existing ties SANCTION.

- Refuse to engage in or expand economic ties.

- Use this code for rejections of mutual economic exchange, such as trade and investment; rejection to provide financial aid (or cancel debt) is coded as CAMEO 1221 instead.

- Refuse to engage in or expand military ties.

- Use this code for rejections of mutual military exchange; rejection to provide military aid is coded as SANCTION instead.

- Refuse to engage in or expand cooperation in judicial matters, including extraditions or other matters pertaining to legal proceedings.

- Refuse to engage in or expand cooperation in intelligence or information sharing.

- Refuse to extend financial, military or humanitarian assistance. Refusals to provide shelter or refuge should also be coded here.

- Refuse to provide peacekeeping forces or other form of military protection;; refusals by adversaries to grant access to peacekeepers.

- Refuse to change leadership or relinquish power.

- Refuse to change a given policy.

- Refuse to provide or respect social, political, economic or other rights and freedoms.

- Refuse to make fundamental political changes, such as moving from one type of a political system to another and reforming political institutions or key laws.

- Reject requests, refuse or decline to ease administrative sanctions, such as censorship, curfew, state of emergency, and martial law.

- Reject requests, refuse, or decline to reduce or stop political protest activities, such as demonstrations and rallies.

- Reject requests, refuse, or decline to release or return persons or property.

- Reject requests, refuse, or decline to reduce or eliminate economic sanctions, boycotts, or embargoes.

- Reject requests, refuse or decline to allow access to international actors such as observers, humanitarian agencies, and peacekeeping forces.

- Reject requests, refuse, or decline to stop fighting or take measures to ease military conflict or tension, including ceasefires, military withdrawals, and demobilization.

- Refuse to meet, discuss, or negotiate, including involvement of mediators or mediation initiatives.

- Reject a proposal or request for a final, comprehensive settlement, peace proposal, or resolution.

- Disobey, challenge, or resist laws or norms.This event category covers both civilian disobedience and official defiance.

- Refuse to assent or formally reject legislative proposal, recommendation, or resolution.

2.12.1   Requires target: No

2.12.2   Supplementary fields:

None?—for the time being we're going to see whether everything can be done with *context* rather than setting up *mode* categories.

2.13 THREATEN

All threats, coercive or forceful warnings with serious potential repercussions. Threats are typically verbal acts. Examples from the CAMEO manual include:

- Threats to reduce or eliminate provision of material assistance—economic, military, humanitarian, and peacekeeping.

- Threaten to restrict normal economic interactions by imposing sanctions, boycotts, or embargoes.

- Threaten to reduce or formally sever ties. Non-force threats to declare independence, resign, withdraw diplomats, reduce or break diplomatic ties, etc. are all coded here.

- Threaten to impose or expand non-force administrative restrictions and penalties not otherwise specified.

- Threaten to impose or expand restrictions on fundamental freedoms, such as freedoms of speech, expression, and assembly.

- Threaten to ban political activities of particular parties or individuals.

- Threatened with imprisonment or other measures of repression.

- Threaten to enforce a deadline beyond which inhabitants of an area are not permitted to be on the streets or in public places.

- Threaten with suspending certain given rights or the whole constitution by imposing state of emergency or military rule.

- Threaten to mobilize or engage in actions of political dissent such as protest demonstrations, hunger strikes, strikes or boycotts, physical obstructions into buildings or areas, and riots.

- Threaten to break-up or withdraw from discussion, negotiation, or meeting.

- Threaten to reduce or stop international intervention by expelling or withdrawing observers, humanitarian agencies, peacekeepers, etc.

- Threats by international agencies to withdraw their involvement as well as threats by host countries to expel such actors are coded here.

- Threaten dissidents with forcible subjugation.

- Threats to imprison as well as to use force to clamp down on opposition activities are coded here.

- Threaten to prevent entry into and/or exit from a territory using military measures.

- Threaten to occupy, seize control of the whole or part of a territory. This event form is typically a verbal act and is distinct from FIGHT, which refers to military occupations that have been or are being carried out.

- Threaten to use violence, including terrorist activities

- Give a final warning, ultimate demand or order, the rejection of which carries the risk of some form of retaliation by the party issuing the ultimatum.

2.13.1 Requires target: No

2.13.2 Supplementary fields

Table 2.2: THREATEN modes

Name	Content
restrict	restrict movement of people or goods, including boycotts, strikes, blockades, and curfews
ban	threaten to ban political activities of particular parties or individuals
arrest	arrest, detain, imprison
relations	threaten to suspend relations, talks
political	ban political activities or restrictions on fundamental freedoms, such as speech, expression, and assembly
expel	expel diplomats, peacekeepers, NGOs
territory	threaten to occupy, seize control of the whole or part of a territory
violence	threaten violence

2.14   PROTEST

All civilian demonstrations and other collective actions carried out as protests against the target actor: Dissent collectively, publicly show negative feelings or opinions; rally, gather to protest a policy, action, or actor(s).

2.14.1   Requires target: No

2.14.2   Supplementary fields:

context: Context (topic) of protest: see Table 2.3

mode: Mode of protest: see Table 2.4

size: number of participants (integer or code)

injured: number injured

eventLoc: Location of event

Table 2.3: PROTEST contexts

Name	Content
election	elections
political	political and constitutional reforms
economic	economy, jobs
food	food, water, subsistence
environment	environmental degradation
discrimination	ethnic discrimination, ethnic issues
religion	religious discrimination, religious issues
education	education
foreign	foreign affairs/relations
war	domestic war, violence, terrorism
rights	human rights, democracy
pro-govt	pro-government
assets	economic resources/assets
other	other
unknown	unknown, not-specified

Adapted from Salehyan and Hendix, *Social Conflict Analysis Database* (SCAD) Version 3.2:
`https://www.strausscenter.org/images/codebooks/SCAD_32_Codebook.pdf`

Observations based on coding CAMEO GSRS

1. Added "political"

2. "ethnic" might be a better term than "discrimination'

3. PTB: Do we want to compare these classifications of modes to those used by Chenoweth and MECC?

Table 2.4: PROTEST modes

Name	Content
demo-org	Organized Demonstration. Distinct, continuous, and largely peaceful action directed toward members of a distinct 'other' group or government authorities. Clear leadership or organization(s) can be identified.
demo-spon	Spontaneous Demonstration. Distinct, continuous, and largely peaceful action directed toward members of a distinct 'other' group or government authorities. Clear leadership or organization cannot be identified.
riot-org	Organized Violent Riot. Distinct, continuous and violent action directed toward members of a distinct 'other' group or government authorities. The participants intend to cause physical injury and/or property damage. Clear leadership or organization(s) can be identified.
riot-spon	Spontaneous Violent Riot. Distinct, continuous and violent action directed toward members of a distinct 'other' group or government authorities. The participants intend to cause physical injury and/or property damage. Clear leadership or organization(s) cannot be identified.
strike-gen	General Strike. Members of an organization or union engage in a total abandonment of workplaces and public facilities.
strike-lim	Limited Strike. Members of an organization or union engage in the abandonment of workplaces in limited sectors or industries.
strike-hun	Hunger Strike (from CAMEO 142x).
boycott	Boycott (from CAMEO 143x).
obstruct	Obstruct passage (from CAMEO 144x).

Adapted from Salehyan and Hendix, *Social Conflict Analysis Database* (SCAD) Version 3.2:
`https://www.strausscenter.org/images/codebooks/SCAD_32_Codebook.pdf`

Observations based on coding CAMEO GSRS

1. At the sentence level, it is usually not going to be possible to differentiate the "org" and "spon" status, so perhaps this should just to reduce to "demo" and "riot"

2. "ethnic" might be a better term than "discrimination'

2.15 CRIME

CRIME events are non-political actions which are considered crimes in the jurisdiction where they occur: Table 2.5 lists the common examples. This category is not intended for the coding of acts of civil disobedience, revolt and other activities which, while criminal from the perspective of the government, are primarily political in nature.

2.15.1 Requires target: No

2.15.2 Potential ambiguities

- There is often a great deal of ambiguity as to whether some activities are criminal or political: for example, confiscatory activities by a weak militarized group with little local support. Usually such distinctions will not be apparent at the sentence—or even article—level and need to be resolved elsewhere in the analysis, for example by the classification of the actor.

2.15.3 Supplementary fields

mode: Mode of crime: see Table 2.5 (PTB: What is the source of these crime modes?)

size: any number; typically monetary amount but, for example, could be number of credit card numbers stolen

eventLoc: Location of event

Table 2.5: CRIME modes

Name	Content
murder	murder
assault	assault
sex-violence	sexual violence
sex-work	illegal provision of sexual services for money or goods
theft	theft, robbery and burglary, including vehicular theft
kidnap	kidnapping and hijacking
narcotics	narcotics, including production, transport and sale
smuggling	smuggling (property)
trafficking	smuggling/trafficking (humans)
trespass	trespass; illegal mining, logging, fishing
arson	arson
vandalism	vandalism and other destruction of property
extortion	extortion
corruption	bribery and other corruption of officials
financial	money laundering, tax evasion, insider-trading, embezzlement
cyber	cyber-crime
war	war crimes

2.16 SANCTION

All reductions in normal, routine, or cooperative relations not otherwise specified. Note that this is not confined to formal "sanctions"—SANCTION was just the best word we could find for WEIS and CAMEO's "REDUCE RELATIONS"

Examples from the CAMEO manual include:

- Curtail, decrease, break, or terminate diplomatic exchange.

- Cancellation of meetings, withdrawal, or expulsion of diplomats and termination of other diplomatic activities

- Reductions or terminations of aid not otherwise specified.

- Decrease or terminate provision of economic, military or humanitarian aid.

- Stop or restrict commercial or other material exchange as a form of protest or punishment.

- Terminate discussions, negotiations. Use this event form to code failed negotiations and walkouts, as well as other disruptions of planned negotiations. Note that the termination can be either unilateral or bi/multi-lateral.

- Terminate mediation activities.

- Terminate the presence of groups or organizations: this covers both expulsions by host authorities and withdrawals by guest groups or organizations, as well as diplomats are withdrawn or expelled. .

- Terminate the deployment or presence of peacekeeping forces, inspectors or other observers.

- Terminate the presence of aid agencies or other non-governmental organizations helping civilians.

2.16.1 Requires target: Yes

2.16.2 Potential ambiguities

- Expulsions or deportations of individuals—typically a legal matter—are coded as COERCE

- Withdrawal of hostile military forces constitutes a form of yielding and is thus coded as YIELD.

2.16.3 Supplementary fields:

None?—for the time being we're going to see whether everything can be done with *context* rather than setting up *mode* categories.

2.17 MOBILIZE

All military or police moves that fall short of the actual use of force. This category is different from ASSAULT and FIGHT, as they refer to uses of force, while military posturing falls short of actual use of force and is typically a demonstration of military capabilities and readiness. MOBILIZE is also distinct from THREAT in that the latter refers merely to threats, is typically verbal, and does not involve any activity that is undertaken to demonstrate military power. Source actors are not necessarily militaries affiliated with states but any organized armed groups. Targets are actors against whom the source mobilizes its military capabilities in a threatening manner if that is clear, but a group may mobilize with no specific target stated.

2.17.1 Requires target: No

2.17.2 Supplementary fields:

Table 2.6: MOBILIZE modes

Name	Content
troops	Mobilize armed personnel or units
weapons	Increase readiness of weapons systems (can occur with a `cyber` context)
police	Mobilize or increase readiness of police or security units

Adapted from CAMEO category 15x

2.18 COERCE

Repression, violence against civilians, or their rights or properties.

2.18.1 Requires target: No

2.18.2 Supplementary fields:

Table 2.7: COERCE modes

Name	Content
confiscate	confiscate property
destroy	destroy property
restrict	impose restrictions on political freedoms or movement
ban	ban individuals or organizations
censor	censor, ban or restrict access to publications
curfew	impose curfew
martial-law	impose state of emergency or martial law
arrest	arrest, detain, or charge with legal action
deport	expel or deport individuals

Adapted from CAMEO category 17x

2.19   ASSAULT

ASSAULT events are deliberate actions which can potentially result in substantial physical harm: Table 2.8 lists the common modes.

2.19.1   Requires target: No

In ASSAULT events where the violence is two-sided, all participants are coded as source actors. In one-sided violence, the perpetrator is coded as the *source* and the victim as the *target*.

2.19.2   Supplementary fields:

Table 2.8: ASSAULT modes

Name	Content
beat	physically assault
torture	torture
execute	judicially-sanctioned execution
sexual	sexual violence
assassinate	targeted assassinations with any weapon
primitive	primitive weapons: fire, edged weapons, rocks, farm implements
firearms	rifles, pistols, light machine guns
explosives	any explosive not incorporated in a heavy weapon: mines, IEDS, car bombs
suicide-attack	individual and vehicular suicide attacks
heavy-weapons	crew-served weapons
other	other modes

Adapted from Political Instability Task Force Atrocities Database:
`http://eventdata.parusanalytics.com/data.dir/atrocities.html`

mode: Mode of violence: see Table 2.8

dead: number killed (integer or code)

injured: number injured (integer or code)

size: used when total casualties are reported, combining dead and wounded

eventLoc: Location of event

2.20   DOCUMENT

DOCUMENT "events" are means of including some internal documentation in the data set while still allowing the file to be read as a single JSON object: this is similar to the DOC event used in the data sets produced by TABARI. Generally these records will only use the *citation*, *publication*, *coder*, *version*, *dateCoded*, and *comment* fields: detailed documentation should be placed in a separate file with a reference, typically, in the *comment* field. However, the specifics of the use of this field are left to the data provider. DOCUMENT events should be skipped when doing statistical analyses of the data.

2.20.1 Requires target: No

2.20.2 Supplementary fields:None

2.21 Context codes that can be used with any category

Table 2.9: General contexts

Name	Content
political	political contexts not covered by any of the more specific categories below
military	military, including military assistance
economic	trade, finance and economic development
diplomatic	diplomacy
resource	territory and natural resources
culture	cultural and educational exchange
disease	disease outbreaks and epidemics
disaster	natural disaster
refugee	refugees and forced migration
legal	national and international law, including human rights
terrorism	terrorism
government	governmental issues other than elections and legislative
election	elections and campaigns
legislative	legislative debate, parliamentary coalition formation
cbrn	chemical, biological, radiation, and nuclear attacks
cyber	cyber attacks and crime
historical	event is historical
hypothetical	event is hypothetical

Chapter 3

Actor and Sector Codes

CAMEO employed a hierarchical actor coding structure based on 3-character coding elements which allowed nearly unlimited complexity and, depending on the exact coding system, could be resolved down to the identity of individual groups or individuals. As with the event codes, typically only the first two or three of these elements were used. ICEWS modified this somewhat, while preserving most of the sub-state differentiations as "sectors"—the terminology we've adopted here over the CAMEO/IDEA "agents"—but also provided a very substantial amount of complexity at the sub-sector level.

As with the events, the PLOVER specification seeks to pare this down to the most commonly used actors and agent/sectors, while retaining the possibility of more specific information. In place of the pages of actor and agent specification found in the CAMEO manual, PLOVER has four rules:

1. The actor code—*source/target*—is either an ISO-3166-alpha-3 code or one of a small number of non-state codes

2. The sector code is a 3-character primary code with one optional secondary code

3. Identifiers for individual persons or organizations are coded in the $*Identifer$ and/or $*Text$ fields

4. The $*Info$ JSON object is used for additional information beyond what can be coded in the sector secondary modifer[1] such as religion, ethnicity, official position, etc

   2-element sector codes?? You're going to regret that, right?

   Yeah, probably. The primary motivation for this is differentiating multiple sub-state militarized groups (REB) which in the current environment are typically distinguished by either religion (Central African Republic), ethnicity (Somalia, South Sudan, Myanmar) or a combination (Iraq, Syria). Allowing a second element in the sector codes allows one to do this without using the more elaborate $*Info$ object. But we're not allowing sector codes to handle something of the complexity of "Iranian-backed Bosnian Serb militia."[2]

---

[1] In other words, one item of information—typically it is religion, ethnicity, or position—can be coded in the tertiary code, but only one: this handles virtually all of the current use-cases we know of.

[2] Such a group existed in the former Yugoslavia in the early 1990s: we're guessing it was rather small.

3.1 Actor codes

For nation-states and other entities for whom an ISO-3166 code[3] exists[4], use the alpha-3 code. Use the codes in Table 3.1 for non-state actors:

Table 3.1: Non-state actor codes

Code	Content
IGO	international governmental organization
NGO	non-governmental organization
ISM	international social movement (See note 1)
IMG	transnational militarized group
MNC	multi-national corporation

Notes:

1. This also includes "international persons" such as Jimmy Carter, Bono, and Angelina Jolie

3.2 Sector Primary Code

Table 3.2: Sector Primary Codes

Code	Frequently used codes
GOV	Government: the executive, governing parties, coalitions partners, executive divisions
JUD	Judiciary: judges, courts
LEG	Legislature: parliaments, assemblies, lawmakers
MIL	Military: troops, soldiers, all state-military personnel/equipment
COP	Police forces, officers, criminal investigative units, protective agencies
OPP	Political opposition: opposition parties, individuals, anti-government activists
PTY	Political parties not identified with government or opposition
REB	Rebels: armed opposition groups or individuals (see Note 1)
PRM	Paramilitary organizations not in opposition to government
SPY	State intelligence services
UAF	Unidentified armed forces ("unknown gunmen")
	Less frequently used codes
CVL	civilians: sometimes used as catch-all for individuals
ELI	clites: former government officials, celebrities, spokespersons for organizations
BUS	usiness: individuals companies, and enterprises, not including MNCs
EDU	educators, schools, students, or organizations dealing with education
MED	endividuals and organizations dealing with health (see Note 2)
LAB	formally or informally organized labor in services or manufacturing
AGR	formally or informally organized agricultural labor; peasants
JRN	journalists, newspapers, radio, television, web sites (see Note 2)

[3]`https://en.wikipedia.org/wiki/ISO_3166-1_alpha-3`
[4]For example, dependencies such as the Åland Islands

REF	refugees and internally displaced persons
REL	religious organizations and institutions
SOC	any organization or movement that is considered part of "civil society" not otherwise covered here
CRM	individual criminals and criminal gangs

Notes:

1. For militarized groups, we are dropping the INS (insurgent) and SEP (separatist) distinctions incorporated into CAMEO during the research phase of ICEWS: these can be resolved on the basis of the group identity and group objectives are frequently ambiguous in any case.

2. This is the one instance where we've altered the meaning of a CAMEO sector code: in CAMEO 'MED' was "media" and 'HLH' was "medical" but no one could remember those.

3.2.1 Sector Secondary Code

We are leaving this open, but for purposes of standardizing, would point the user to the following resources we have found useful

CAMEO Religious Codes: This is a very large hierarchically-organized scheme for coding virtually all religions and religious-movements that could be identified ca. 2010. LaTeX and PDF versions of this are available at `https://github.com/openeventdata/PLOVER`

CAMEO Ethnic Codes: This was CAMEO's effort to provide a uniform set of ethnicity codes based on two primary sources: (1) the International Organization for Standardization's (ISO) Codes for the Representation of Names of Languages (`ISO-639.2;http://www.loc.gov/standards/iso639-2/`) and (2) the Ethnic Power Relations (EPR) dataset 3.1 (`http://www.epr.ucla.edu/`). LaTeXand PDF versions of this are available at `https://github.com/openeventdata/PLOVER`.

ICEWS Sectors: The file `icews.sectors.20140112.csv` provides three-level hierarchy of sectors which could be used as an alternative to the specification here; `icews.agents.20140112.csv` associates these with common names or noun-phrases found in news reports: `https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/28118`

PETRARCH Agents: PETRARCH project agents file, which was assembled from earlier KEDS project TABARI files, WordNet (`https://wordnet.princeton.edu/`), CIA World Leaders and an assortment of other sources. `https://github.com/openeventdata/petrarch2/blob/master/petrarch2/data/dictionaries/Phoenix.agents.txt`

3.2.2 Compound and Reciprocal Actors

Most CAMEO-based coders dealt with compound actors—"The United States and France accused Russia..."—by generating multiple events: this example would generate two events of the form

```
USA  RUS  112
FRA  RUS  112
```

This approach, however, gets very problematic in the not-uncommon situation where an alliance is involved and is expanded to all of its constituent members: a single reference to the G20 expands to at least twenty events, and a *meeting* of the G20, generating reciprocal events, expands to 380 events, which is one of the reasons "consult" events are so frequent in CAMEO-coded data.

In PLOVER, compound actors generate a single source or target, but with multiple members: the source and target are a list of actors rather than a single actor. Depending on the application, a user might expand this to multiple events with single actor codes following the earlier conventions, but the initial coding uses the list.

PLOVER also uses actor lists to deal with reciprocal events. In CAMEO, a meeting "President Obama met with Japanese Prime Minister Abe at the White House"" generated two events

```
JAP  USA  042
USA  JAP  043
```

where 042 and 043 are the CAMEO codes for "visit" and "host" respectively.[5] While the PLOVER CONSULT mode provides for a host/visit distinction, this is not required. In such instances, all of the actors are considered as the source and no target is included.

In ASSAULT events, reciprocal violence—as distinct from one-sided violence—is handled in a similar fashion, with both parties as *source* actors: this applies in any event where both sides are using force, even if one side "started it"[6] One-sided violence, such as assassinations or police firing on demonstrators, will have the perpetrator as the *source* and the victims as *target*.

---

[5]This example optimistically assumes the coding system was clever enough to recognize the significance of the phrase "at the White House."

[6]An assessment often as not contested anyway.

Chapter 4

JSON Data Fields

In keeping with OEDA's founding aversion to any efforts at creating "one data set to rule them all," almost all of the fields listed below are optional: the objective of this part of the specification is mostly to provide a standard set of field names to simplify the merging and reuse of datasets. Despite the apparent complexity here, note that the only required field we have added to "event data classic" is the *id* identifier, so the simplest form of an event record would look like

```
{
    "id" : "PHOXv1-20160724-0042",
    "date" : 2016-07-24",
    "source" : [{"code":"USA"}],
    "target" : [{"code":"CAN"}],
    "event" : "AGREE"
}
```

Except in the small number of cases where a standard format is specified, the content of the field is left open, and in particular "number" should be interpreted as "number or code": for example instead of providing the number of individuals killed, a dataset might use a set of categories giving ranges. Similarly, categories such as *context* can take multiple values: typically your end users would probably be happiest if these were formatted using a JSON "array" structure—comma-delimited in square brackets—but this is not assumed. Responsibility for handling these details is left to the data provider and users.

Providers should feel free to include named fields beyond those provided hence—hence the absence of any generic field names like *extra*1, *extra*2 in the standard—but if you are coding or extracting information that corresponds to one of the existing fields, please use that name.

Table 4.1: PLOVER JSON

Name	Content	Note	Required?
id	unique identifier	1	Y
date	date in YYYY-MM-DD format		Y
time	ISO 8601-formatted time	2	N
enddate	date in YYYY-MM-DD format		Y
endtime	ISO 8601-formatted time	2	N
source	list of actor objects		Y
target	list of actor objects		N
event	event category		Y
eventLoc	location object for event		N
eventText	text of event		N
quadCode	1, 2, 3 or 4		N
eventScale	floating point scale value		N
mode	mode category	3	N
context	context category	3	N
dead	number killed		N
injured	number injured		N
size	number: depends on context		N
link	link identifier	4	N
text	text from which the record was coded		N
citation	bibliographic citation or database identifier for text		N
url	URL for text		N
language	language of text (ISO 639-1 two-letter codes)		N
publication	name of text publisher		N
license	license covering text		N
copyright	copyright covering text		N
textInfo	textInfo object for text		N
coder	coder identification		N
version	version of data set		N
codebook	reference for the codebook used to code the text		N
dateCoded	date of coding		N
comment	any text		N

Notes:

1. The identifier should be unique within the data set; it is the responsibility of the user to reconcile identifiers across data sets

2. ISO 8601 allows a number of different formats for times depending on the level of detail. A string of the form `date` + 'T' + `time` should yield an ISO-8601 datetime.

3. *mode* and *context* fields can have multiple entries; they do not need to resolve to a single value.

4. This can be used to create a common reference across multiple related events, demonstrations in multiple locations organized by the same group.

5. PTB Question: Should there also be lat and lon for any geolocation of the eventLoc? PAS: eventLoc has that inside the location object

Table 4.2: Information object for actors

Name	Content
code	3-char actor code
sector	3- or 6-char source sector
identifer	unique identifier for source [see Note 1]
actorLoc	location object
actorText	extracted text for source
religion	religion (code or text)
ethnicity	ethnicity (code or text)
office	office or official position (code or text)
gender	gender (code or text)
age	integer

Notes:

1. These fields would be used to resolve the name of an actor that occurs in multiple forms—for example "Islamic State", "IS", "ISIS", "Daesh"—into a single form or code (for example the organization number in the TORG typology). PTB: Need to resolve to a link for TORG.

Table 4.3: Information object for text

Name	Content
sequence	sequence number of sentence
start	character offset for start of text
end	character offset for end of text
markup	text with markup, for example parsing

Table 4.4: Location object: identical to http://download.geonames.org/export/dump/readme.txt

Name	Content
geonameid	integer id of record in geonames database
name	name of geographical point (utf8)
asciiname	name of geographical point in plain ascii characters
alternatenames	alternatenames, comma separated, ascii names automatically transliterated
latitude	latitude in decimal degrees
longitude	longitude in decimal degrees
feature class	see http://www.geonames.org/export/codes.html
feature code	see http://www.geonames.org/export/codes.html
country code	ISO-3166 2-letter country code, 2 characters [see Note 1]
cc3	ISO-3166 3-letter country code [see Note 1]
cc2	alternate country codes, comma separated, ISO-3166 2-letter country code,
admin1 code	fipscode (subject to change to iso code)
admin2 code	code for the second administrative division, a county in the US
admin3 code	code for third level administrative division
admin4 code	code for fourth level administrative division
population	bigint (8 byte int)
elevation	in meters, integer
dem	digital elevation model: see geonames documentation for details/ciat.
timezone	the iana timezone id

Notes:

1. Geonames, alas, uses ISO-3166-alpha-2 rather than the more mnemonic alpha-3 codes used in most event data work. For purposes of compatibility, we're suggesting retaining this in the "country code" field but adding a "cc3" field (not found in geonames) for alpha-3 codes.

2. See https://github.com/openeventdata/PLOVER/issues/3

Chapter 5

CAMEO vs. PLOVER

5.1   CAMEO to PLOVER translation

Table 5.1: PLOVER equivalents to CAMEO cue categories

CAMEO code	CAMEO text	PLOVER category
01	MAKE PUBLIC STATEMENT	dropped
02	APPEAL	dropped
03	EXPRESS INTENT TO COOPERATE	AGREE
04	CONSULT	CONSULT
05	ENGAGE IN DIPLOMATIC COOPERATION	SUPPORT
06	ENGAGE IN MATERIAL COOPERATION	COOPERATE
07	PROVIDE AID	AID
08	YIELD (081 to 083)	CONCEDE
08	YIELD (084 to 087)	RETREAT
09	INVESTIGATE	INVESTIGATE
10	DEMAND	DEMAND
11	DISAPPROVE	DISAPPROVE
12	REJECT	REJECT
13	THREATEN	THREATEN
14	PROTEST	PROTEST
15	EXHIBIT FORCE POSTURE	MOBILIZE
16	REDUCE RELATIONS	SANCTION
17	COERCE	COERCE
18	ASSAULT	ASSAULT
19	FIGHT	ASSAULT
20	USE UNCONVENTIONAL MASS VIOLENCE	FIGHT (see Note 1)
–	no CAMEO equivalent	CRIME

Notes:

1. For unconventional weapons, the *mode* in the FIGHT record would be set to "Unconventional." In PLOVER, mass expulsions, killings and ethnic cleansing (CAMEO 201, 202, and 203) are a pattern of events, not a single event.

2. Generally, everything at the 3- and 4-digit level should simply be reduced to the 2-digit cue category and converted accordingly. Depending on your specific application, you might want to make some exceptions to this—for example a CAMEO "015: Acknowledge or claim responsibility" might be considered AGREE and a CAMEO "016: Deny responsibility" might be considered REJECT—but we are not making general recommendations on this. Except to suggest that for the benefit of those trying to replicate your work, you carefully document any such decisions.

5.2 PLOVER quad categories

Table 5.2: Quad categories in PLOVER

Quad category	PLOVER categories	Numeric
Verbal cooperation	AGREE, CONSULT, SUPPORT, CONCEDE	1
Material cooperation	COOPERATE, AID, RETREAT, INVESTIGATE	2
Verbal conflict	DEMAND, DISAPPROVE, REJECT, SANCTION, THREATEN	3
Material conflict	PROTEST, CRIME, MOBILIZE, COERCE, ASSAULT	4

Chapter 6

PLOVER Extensions

One of the objectives of PLOVER is to encourage and provide guidance to researchers wanting to create custom ontologies for event data sets tailored to answering their research questions. PLOVER and this manual try to encourage it in three ways:

1. Making PLOVER much more general than CAMEO to encourage researchers with specific needs to create their own ontologies and event data sets.

2. Making PLOVER less all-encompassing than CAMEO, letting researchers know that custom ontologies do not need to be as complicated and complete as CAMEO.

3. By providing examples and guidance on how to create ontologies.

THIS SECTION IS STILL A WORK IN PROGRESS.

Bibliography

Azar, E. E.
  1980. The conflict and peace data bank (COPDAB) project. *Journal of Conflict Resolution*, 24:143–152.

Azar, E. E.
  1982. *The Codebook of the Conflict and Peace Data Bank (COPDAB)*. College Park, MD: Center for International Development, University of Maryland.

Azar, E. E. and T. Sloan
  1975. *Dimensions of Interaction*. Pittsburgh: University Center for International Studies, University of Pittsburgh.

Bond, D., J. Bond, C. Oh, J. C. Jenkins, and C. L. Taylor
  2003. Integrated data for events analysis (IDEA): An event typology for automated events data development. *Journal of Peace Research*, 40(6):733–745.

McClelland, C. A.
  1967. World-event-interaction-survey: A research project on the theory and measurement of international interaction and transaction. University of Southern California.

McClelland, C. A.
  1976. *World Event/Interaction Survey Codebook (ICPSR 5211)*. Ann Arbor: Inter-University Consortium for Political and Social Research.

McClelland, C. A.
  1983. Let the user beware. *International Studies Quarterly*, 27(2):169–177.

O'Brien, S. P.
  2010. Crisis early warning and decision support: Contemporary approaches and thoughts on future research. *International Studies Review*, 12(1):87–104.

Schrodt, P. A.
  2006. Twenty years of the Kansas event data system project. *The Political Methodologist*, 14(1):2–8.

Schrodt, P. A., D. J. Gerner, and Ö. Yilmaz
  2009. Conflict and mediation event observations (CAMEO): An event data framework for a post Cold War world. In *International Conflict Mediation: New Approaches and Findings*, J. Bercovitch and S. Gartner, eds. New York: Routledge.