

# IBM Research Report

## KLUE Annotation Guidelines Version 2.0

**Ding-Jung (Benjamin) Han**  
IBM Research Division  
Thomas J. Watson Research Center  
P.O. Box 218  
Yorktown Heights, NY 10598



# KLUE Annotation Guidelines

## Version 2.0

Ding-Jung (Benjamin) Han  
dbhan@us.ibm.com

September 2, 2010

## 1 Introduction

This report documents the KLUE annotation scheme: KLUE stands for Knowledge from Language Understanding and Extraction, and it has been the guiding principles behind our internal annotation effort to provide training and testing data for building systems of mention detection, coreference resolution and relation detection.

The previous incarnation of KLUE is KDD annotation guidelines (Knowledge Discovery and Data-Mining). Compared to KDD, KLUE has the following changes, motivated by the performance studies of our systems and by the needs of different applications:

1. Entity types have been consolidated to reduce ambiguity. For example, COMPANYROLE has been eliminated for nominals such as “software makers” - they are now marked as ORGANIZATION and can be coref’ed with a company’s name.
2. Many new relation types have been introduced; e.g. familial relations have seen finer-grained additions such as siblingOf and parentOf.
3. Roles are introduced for mentions to capture metonyms. E.g., the mention “France” in “they went to France” is marked with entity type GPE but with role LOCATION.
4. Subtypes are introduced for entities to capture finer-grained typographical information. E.g., “army” is marked with entity type ORGANIZATION with subtype MILITARY.
5. Anchor words are introduced for relation mentions to capture the “trigger words” for the corresponding relation mention. E.g., in “he was born in 1969”, the anchor word of the relation mention “he” bornOn “1969” is “born”.
6. All cross-type coreference restrictions have been lifted - in KLUE we no longer require only mentions of the same entity type can co-refer to each other.
7. For temporal expressions, we always mark the maximal possible extent instead of breaking a complete temporal expression into multiple sub-expressions. E.g., “Wednesday at 5pm” is marked as one single TIME mention instead of being marked as one DATE mention and one TIME mention.

The rest of this report is organized as follows. In Section 2 we first introduce mentions and their associated annotations (mention types, entity types and roles). The event-related entity types are deferred to Section 3 so that the general instructions on annotating them can be discussed in one place. Next in Section 4 the guidelines on annotating entities (or “coreference chains”) are described. Finally Section 5 describes the relation-related annotations (relation types, tense, modality, specificity and anchor words).

## 2 Mentions

The most basic task in KLUE is to annotate mentions and their associated information. By mention we mean one or several actual words referring to a certain thing of interest. E.g., the words "IBM" and "United States" are both mentions. Several notes on marking a mention:

1. Do not include a punctuation if it is not part of the mention; e.g., "... said the [reporter]." However if a punctuation is part of the mention, do include it in the mention; e.g., "the [U.S.S.R.] era", "the [Sarbanes-Oxley Act]".
2. In some languages a mention can occupy just part of a word: in those cases only mark the portion of the word as the mention. E.g., in Arabic possessive pronouns combine with nouns, so you need to mark them as two separate mentions.
3. In some languages a word and its "marker" can be separated: for example in Arabic the tense marker and the verb can be separated. Mark only the main word in those cases.
4. Do not mark words that serve as search tags - these often appear at the end of a news article, with no context.
5. Do not mark over-generic words as a mention unless they can be (1) coref'ed with a more specific mention, or (2) has a relation with another mention; examples:
  - FOOD: do not mark these - "we had lunch", "there is some food left in the fridge"; but do mark these: "I brought an [orange] today", "he had a [hamburger] for lunch", "we had [Thai food]/FOOD as lunch, and [it]/FOOD was good".
  - ORGAN: do not mark these - "organ transplant is still a risky business", "he wants to donate his organ when he dies"; mark these - "the major [organs]/ORGAN were removed, including his [kidney]/ORGAN" (relation is partOfMany).
  - PRODUCT: do not mark these - "many new products were introduced last year"; do mark these - "one of the new [products]/PRODUCT introduced was [LEGO] [Brainstorm]/PRODUCT" (relation is partOfMany).
  - SUBSTANCE: do not mark these - "chemicals were used to kill germs", "substance abuse is getting worse these days"; mark these - "they used [metals] as the casing".
  - TITLEWORK: do not mark these - "the library bought many new books this year"; do mark these - "the library bought many new [books]/TITLEWORK this year, including [Narnia]/TITLEWORK" (relation is partOfMany).
  - WEAPON: do not mark these - "this is one major breakthrough in weapon development"; do mark these - "the [murder]/EVENT\_VIOLENCE [weapon]/WEAPON was found... [it]/WEAPON has blood stain when it's found", "they found [his]/PERSON [weapons]/WEAPON in the basement" (relation is "his ownerOf weapons").

Below we describe the three pieces of information that comes with each mention: entity type, role and mention type.

- **Entity types**

We use the word entity to refer to the "thing" that a mention refers to. Since, by definition, mentions are coreferential if (and only if) they refer to the same real-world entity, we may use the term "entity" interchangeably with the term "coref(erence) chain" (for coreference annotation see (see Section 4 for the guidelines on coreference annotation).

An entity type indicates what type of an entity a mention refers to, without considering the context. E.g., the word "IBM" refers to an entity of type ORGANIZATION while the words "United States" refers to an entity of type GPE. The rest of this document will describe the details of each entity type (see Section 2.3 for the list of entity types and their roles).

It is important to emphasize the context-free nature of entity types: for every mention with

the same spelling, we must try as much as possible to annotate it with the same entity type. We leave the "context-sensitive" annotation to "roles" (see below).

- **Roles**

Since an entity type only records the context-free meaning of a mention, when the mention is used in place of another kind of entity, we need to annotate this other entity type as the role of the mention (this is called metonymy). E.g., although the entity type of "IBM" is ORGANIZATION, when used in the sentence "I went to IBM today" it should be annotated with role FACILITY. Another example: although the entity type of "United States" is GPE, when used in the sentence "the United States today announced..." its role is ORGANIZATION (the US government).

Several things to note about roles:

- A role is the context-sensitive meaning of a mention (recall that its entity type is its context-free meaning).
- A role itself must be an entity type, and it can also have a sub-type. E.g., the role of "United States" in the sentence "United States today announced..." is ORGANIZATION.GOVERNMENT (see Section 2.1 for explanation on entity sub-types).
- For every entity type there is a fixed set of applicable roles. Therefore we will enumerate them with each entity type in Section 2.3.
- The default role of a mention is its entity type. The default is applicable when the context-sensitive meaning of the mention agrees with its context-free meaning, or when ambiguity prevents us from choosing a sensible role.
- Notation: in the rest of the document we will write a role with its entity type in this form: "IBM" is ORGANIZATION/FACILITY in the sentence "I went to IBM today" (note the slash between the entity type and the role).

- **Role OCCUPATION**

The role OCCUPATION is the only exception to the rule that says every role must be an entity type - OCCUPATION exists only as a role for entity type PERSON, PERSONPEOPLE and PEOPLE. It should be used for mentions of a job that someone can train for (e.g. lawyer, general, Prime Minister) instead of being born to (Emir, King = SALUTATION/PERSON) or falling into (addict, hero = PERSON), and the person has to gain monetary reward from doing the job through actual work (e.g., "co-founder" of a company is an occupation, but "shareholder" is not because there is no real work involved). Examples:

- “Russian [President]/PERSON/OCCUPATION Boris Yeltsin”
- “C. Everett Koop , the former U.S. [surgeon general]/PERSON/OCCUPATION”
- “the [lawyers]/PEOPLE/OCCUPATION”

- **Mention types**

Each mention can be a name (e.g., "IBM"), a nominal ("company"), a pronoun ("it"), or none of the above ("three"). We annotate this fact via the mention type of a mention. Corresponding to the four possibilities mentioned above, the possible mention types are NAM, NOM, PRO or NONE.

For a GPE mention, if its role is PERSON/PEOPLE/PERSONPEOPLE, the mention type should be NOM; if its role is anything else, the mention type should be NAM. E.g., "[Syrian] people" is NOM, "[Syrian] government" is NAM.

A mention of the following entity types should always have a NONE mention type:

- DATE
- DURATION
- EMAIL
- MEASURE
- MONEY
- PHONE
- TICKER
- TIME

## 2.1 Entity Subtypes

Associated with each entity (coref chain) is an entity sub-type, which gives a finer-grained typological information for the entity type. E.g., the entity type of "IBM" is ORGANIZATION, and its entity sub-type is COMMERCIAL. Just like entity types, entity sub-types are also context-free.

Notation-wise we will write a sub-type with its entity type in this form: "IBM" has entity type/sub-type ORGANIZATION.COMMERCIAL (note the dot separating the type and the sub-type).

Do not confuse entity sub-types with roles! A simple test to see whether B is a role or an entity sub-type of A is: if B is a sub-type of A, we can always say "B is an A", but if B is a role of A, we cannot say the same. E.g., a commercial organization is an organization, but a facility is not always an organization (recall one of the roles of ORGANIZATION is FACILITY). Also, entity sub-types are per-entity attributes, while roles are per-mention attributes.

The default sub-type "OTHER" is always available if a sub-type can be identified but is missing from the inventory of this guidelines, therefore "OTHER" will not be listed in the individual sections of the entity types.

For entity types that are not DATE or TIME and have at least two sub-types defined other than "OTHER", when facing ambiguous cases the sub-type "UNSPECIFIED" should be used. The "UNSPECIFIED" sub-type is explicitly added in the individual sections of the entity types.

Although annotating the entity sub-types should be part of the task of identifying entities/coref chains (see Section 4), we will describe the available sub-types per entity type in Section 2.3 for convenience.

## 2.2 General Instructions

### 2.2.1 Modifiers

Modifiers are words modifying a noun, i.e., pre-nominal words ("phone" in "phone booth") and prepositional phrases ("New York" in "City of New York"). We adopt the following procedure to annotate them:

1. If the entire phrase is a well-known term, mark the phrase as a single mention; e.g., "[Gulf war]/EVENT\_VIOLENCE".
2. If words refer to the same entity, mark them as one mention; e.g., "[American people]/GPE/PEOPLE", "[filing suit]/EVENT\_LEGAL".
3. Otherwise, if we have an appropriate entity type for the modifier, and we can find a meaningful relation (not just coref) between the modifier and the noun, annotate the modifier separately from the noun being modified. Examples:
  - "[Jewish] [organization]" = PEOPLE ORGANIZATION and the relation is "organization affiliatedWith Jewish".
  - "[Christian] [organization]" = PEOPLE ORGANIZATION and the relation is "organization affiliatedWith Christian".
4. Otherwise, include the modifier along with the noun as one mention if you feel the modifier is essential for the identification of the mention or the modifier and the noun refer to the same thing. Examples:

- "[phone booth]" because we don't have any type for "phone", and "phone booth" is different enough from "booth" so we mark both words as one mention.
  - "[oil company]" because we cannot find a useful relation between "oil" and "company", although "oil" could be marked as SUBSTANCE.
  - "[city of New York]" (one mention) because there is no meaningful relation between "city" and "New York" (New York is not located at the city, vice versa), and they refer to the same thing.
5. For demonstratives such as "this", "that", and "those": do not annotate them if they co-occur with the nouns being modified; e.g., do not mark "those" in "those people", but do mark "those" in "those affected are ...".
6. For GPE-like modifiers (e.g., US, American, French, New York etc):
- Always use entity type GPE, but pick the role that will enable you to relate the modifier and the noun via a meaningful relation.
  - Examples (the meaningful relation between the two words is added in the parentheses):
    - "[American] [soldiers]" = GPE.COUNTRY/ORGANIZATION PEOPLE (employedBy).
    - "[American] [actor]" = GPE/PEOPLE PERSON (partOfMany).
    - "[American] [President]": GPE/ORGANIZATION.GOVERNMENT PERSON (managerOf); related: "[mayor] of the [city]" = PERSON GPE/ORGANIZATION.GOVERNMENT (managerOf).
    - "[Republican] [McCain]" = ORGANIZATION.POLITICAL PERSON (memberOf).
    - "[US] [President]": GPE/ORGANIZATION.GOVERNMENT PERSON (managerOf).
    - "[US] [Senator]": GPE/ORGANIZATION.GOVERNMENT PERSON (employedBy).
    - "[American] [borders]" = GPE/LOCATION LOCATION (locatedAt).
    - "[American] [facility]": GPE/LOCATION FACILITY (locatedAt).
    - "[American] [city]": GPE/LOCATION GPE/LOCATION (locatedAt).
    - "[European] [city]" GPE.AREA GPE/LOCATION (locatedAt).
    - "[American] [company]": GPE/LOCATION ORGANIZATION.COMMERCIAL (basedIn).
    - "[US] [Senate]": GPE/ORGANIZATION.GOVERNMENT ORGANIZATION.GOVERNMENT (partOf); related: "[police] of the [city]" = ORGANIZATION.GOVERNMENT ORGANIZATION.GOVERNMENT (partOf).
    - "[American] [cars]": GPE/LOCATION VEHICLE (no relation).
    - "[Israeli] [tanks]": GPE WEAPON (ownerOf).
    - "[Israeli] [attacks]": GPE EVENT\_VIOLENCE (agentOf).
    - "[US Government]": GPE/ORGANIZATION.GOVERNMENT (one mention only because the two words refer to the same thing).
    - "[Russian Government]": GPE/ORGANIZATION.GOVERNMENT (one mention only because the two words refer to the same thing).
    - "[Chinese] [Authorities]": GPE/ORGANIZATION.GOVERNMENT ORGANIZATION (partOf; "authorities" is not necessarily the entire government); related: "[authorities] of the [city]" = ORGANIZATION GPE\_ORGANIZATION (partOf).
    - "[US People]": GPE/PEOPLE (one mention only because the two words refer to the same thing).
    - "[American People]": GPE/PEOPLE (one mention only because the two words refer to the same thing).

Modifiers in conjunctive phrases: don't mark the modifiers if they are used in a conjunctive phrase; examples:

- In "Security and medical officials" mark only "officials" because (a) we have no way to mark "security" with "officials" together and (b) there are no applicable entity types for "security" and "medical".
- We mark "[kidnaping]/EVENT\_CRIME and [murder]/EVENT\_CRIME [trial]/EVENT\_LEGAL" because we do have applicable entity types for the conjunctive modifiers.

## 2.2.2 Pronouns

Pronouns are words such as "it", "they", "she", "where" and they almost always are coref'ed with another mention. They are highly context-sensitive and are one major exception to the principal of "sticking to the same entity type for the same spelling" described earlier.

A pronoun's entity type or role must be equal to the entity type or the role of the mention it is coref'ed with. E.g., "The [Cubans]/GPE/PEOPLE..., [they]/GPE/PEOPLE ... The [people]/PEOPLE" where the three nouns are in the same coref chain - "they" has a role equivalent to the entity type of "the Cubans" and the entity type of "people".

Below are the two classes of pronouns and some examples illustrating this principal (these are English examples but please draw conclusions in your respective language):

- Pronouns such as "it", "they", "he", "that" (as in "that was broken"), "there", "here" etc; examples: "[Mary]/PERSON didn't make it to work yesterday because [she]/PERSON was sick", "the [government]/GPE/ORGANIZATION announced [it]/GPE/ORGANIZATION was banning Hezbollah's military wing".
- Pronouns such as "where", "when", "what", "who", "which", "that" (as in "the car that was broken"), "whose" etc; examples: "US was banning [Hezbollah]/ORGANIZATION, [whose]/ORGANIZATION military wing...", "he visited an [American]/GPE/PERSON [who]/GPE/PERSON moved there long time ago".

## 2.2.3 Case (Capitalization)

Some text comes with case, i.e., some words are capitalized while others are not. Follow the procedure below to deal with case if it is available to your language (e.g., case is not available in Farsi)

- If all words in a name are capitalized, mark them as one mention unless they naturally form multiple names (see below). This is probably the only place where a little bit Googling could help (exception to the general rule that you should not use world knowledge). E.g., "[City of Westminster Magistrate's Court]", "[Swiss Central Bank]", "[Court of Appeals for the 11th Circuit]", "" etc.
- However if not all words are capitalized, you can opt to mark them as one or several mentions. E.g., in "former [Yugoslav] [republic of Macedonia]" and "[Bagram] [airbase]" we created more than one mention, but in "[JFK airport]" we elect to mark both words as one mention instead of PERSON LOCATION, since there is no valid relation to relate the two mentions; we also mark "Gulf war" as one mention since it is a well-known historic event, so there is no need to mark it as LOCATION EVENT\_VIOLENCE and then say the event is locatedAt that location.
- If capitalized words naturally form more than one name, mark them separately; e.g., mark "[Toyota] [Land Cruiser]" as ORGANIZATION.COMMERCIAL VEHICLE (instead of marking both words as one mention).
- If all letters are all-capitalized, or the first letter of a sentence is capitalized for aesthetic reasons, they should be treated as if they are not capitalized; e.g., "JFK AIRPORT" should be treated as "jfk airport".

## 2.2.4 Temporal Expressions

Temporal expressions are phrases indicating time/duration, and they are marked as DATE, DURATION or TIME; e.g., "Monday morning", "Aug 4, 1949", "an hour".

A couple notes on marking the extent:

- A temporal expression must have at least one time unit word in it. E.g., "the last time" is not considered a temporal expression.
- Words such as "around", "about", "nearly" should be included. E.g., "[around 7pm yesterday]", "nearly 2 hours".

- Comparative words such as "longer than" should also be included. E.g., "[longer than 4 days]", "[no less than 3 hours]".
- If omission of a word can change the meaning of the temporal expression, include that word. E.g., "I went there [two days after]"; "[the night] it happened" (definite article indicates it is a specific night).

**Extent and Entity Type of Temporal Expressions** When marking a temporal expression, its maximal extent should be marked; e.g., "[Monday morning]", "[3 in the afternoon last Monday]", "[2 to 3 weeks]", "[3 or 4 hours]" should all be marked as single mentions. When deciding whether a temporal expression is a date or a time, consider the finest granularity it indicates; e.g., both "[Monday morning]" and "[3 in the afternoon last Monday]" should be a TIME, while "the second Sunday of May" is a DATE.

**Event-dependent Temporal Expressions** Some DATE and TIME are meant to be interpreted with an event; e.g., in "two days after the first explosion" the entire phrase denotes a date that is two days after the time of an event "explosion". In such cases we mark the maximal extent of the temporal expression ("two days after" as DATE) and the event ("explosion" as EVENT\_VIOLENCE) separately, and then use temporal relation before to relate the two (so "explosion" before "two days after").

If there are multiple events mentioned in a sentence together with a temporal expression, mark the relations between the events and the temporal expression first. If such a relation cannot be found for an event, then we should attempt to relate that event to another event using a temporal relation. Examples:

- In "The plane [crashed]/E1 [5 minutes after]/E2 the [explosion]/E3", we have E2 = TIME, E2 timeOf E1 and E3 before E2 (we don't mark E3 before E1 since we link every event to the TIME).
- In "the [incident]/E1 [yesterday]/E2, in addition to the previous [attacks]/E3 in the region, signifies the shift of the focus of the organization", we have E1 = EVENT\_VIOLENCE, E2 = DATE, E3 = EVENT\_VIOLENCE, E2 timeOf E1 but E3 before E1; we cannot say E3 before E2 (although ideally we'd want to link every event to a temporal expression) because there might also be an attack yesterday, just before "the incident".
- In "she was [born]/E1 [a week after]/E2 the [earthquake]/E3", mark E1 = EVENT, E2 = DATE, E3 = EVENT\_DISASTER, E2 timeOf E1, and E3 before E2 (note we don't have a more specific entity sub-type for E1).
- In "she was [born]/E1 [a week]/E2 after\_E2 he was [born]/E3", because both E1 and E2 do not have an applicable event sub-type, and assuming neither event is the topic of the document, we don't mark the events, but we do want to mark E2 as DURATION.

## 2.3 List of Entity Types and Their Roles

In this section we will detail each entity type present in KLUE.

### 2.3.1 AGE

Used to classify how many years old someone or something (including an EVENT) is. It must contain numbers, and numbers must be followed by years old or allude to such. Examples: "[50 years old]", "on his [75th] birthday", "Dan Smith, [32]", "[39-year-old]/AGE/DURATION [occupation]/EVENT\_VIOLENCE of the West Bank". Note when it is used with an event, the role should be DURATION.

1. Words like "around", "about" should be included.



- Comparatives such as "older than" should be included (but remember only when numbers present do we mark the expression as AGE; e.g., "she is older than him" should not be marked), and also do not mark expressions such as "she is 5 years older than Tom" (but "she is older than 50" is okay).
- If the age expression is used to refer to the person, mark it AGE but assign the role PERSON. E.g., "the [6-year-old]/AGE/PERSON fell off the bike?".

### Roles

- PERSON: E.g., "the [6-year-old]/AGE/PERSON fell off the bike".
- DURATION: E.g., "[39-year-old]/AGE/DURATION [occupation]/EVENT\_VIOLENCE of the West Bank".

**Sub-types** There is no sub-type for this entity type.

### 2.3.2 ANIMAL

Used to classify animals, real or fictional. Ex: "[Lions]", "[tigers]" "[bear]", "[Smokey the Bear]", "Phyllis's [cat] [Puck]".

Do not mark inside an idiom, e.g., "lipstick on a pig".

### Roles

- PEOPLE: E.g., "scientists who become their own [guinea pigs]"
- PERSON: E.g., "he is such a [pig]."

**Sub-types** There is no sub-type for this entity type.

### 2.3.3 AWARD

Used to classify awards; e.g., Nobel Peace Prize, Emmy, Oscar, featherweight title, etc. Note that for money used as award mark it as MONEY/AWARD.

### Roles

- PERSON: E.g., "the [MVP] said..."
- EVENT\_PERFORMANCE: E.g., "[Oscar] will be held on ..."

**Sub-types** There is no sub-type for this entity type.

### 2.3.4 CARDINAL

Used to classify numerals that provide a count or quantity that does not fall under a unit of measurement.

- Whole numbers, decimals and fractions are included. E.g., "[half]", "[hundreds]", "[one-third]", "[four]", "[4534]".
- Words like "both", "few", "many" and "several" are included if they are applied to count nouns; e.g., "all of the water" should NOT be marked.
- Words like "once", "twice" etc should also be marked as CARDINAL.
- Words like "around", "about" and "approximately" etc are also included.
- Include the comparative words; e.g., "[more than 3000]".

6. Numbers in addresses should also be marked as CARDINAL; e.g., "[18] Main Street".
7. Phone and fax numbers should not be marked as CARDINAL (use PHONE instead).
8. Cardinal expressions that refer to another type of entities should still be marked as CARDINAL, but an appropriate role must be chosen for them. E.g., "[two]/CARDINAL/PEOPLE of the people failed passing the test".

**Roles** Every entity type, except for AGE / CARDINAL / DATE / DURATION / ORDINAL / PERCENT / TIME, can be a role for CARDINAL. E.g., "[one]/CARDINAL/PERSON of the five people", "[two]/CARDINAL/VEHICLE of the cars".

**Sub-types** There is no sub-type for this entity type.

### 2.3.5 DATE

See Section 2.2.4 for the general instructions on all temporal expressions.

Used to classify a date expression. E.g., "[November 2, 2001]", "[60s]", "[mid-1940's]", "[Feb 2]", "[Last week]", "[last year]", "[2 years ago]", "[January]", "[mid-April]", "[Monday]", "[yesterday]", "[tomorrow]", "[Spring]", "[medieval] France".

1. As mentioned here, we should mark the maximal extent for a date expression; e.g., "[Nov 5, 6, and 7]" should be marked as a single mention.
2. Do not mark figurative speech such as "today's youth". Holidays such as "Christmas" and "New Year's Day" should also be marked as DATE.
3. See Section 2.2.4 if a date depends on an event ("two days after the first explosion"). Note "the day after tomorrow" is not an event-dependent expression - it is a single DATE.
4. Financial quarter: since every company defines its own financial calendar, we don't mark it; however, if the month is explicitly mentioned, e.g., "September quarter", we mark the entire span as one DATE.

**Roles** There is no role for this entity type.

**Sub-types**

- STARTING: used to classify the DATE that serves as a starting time-point. E.g., "the game started on [Saturday]."
- ENDING: used to classify the DATE that serves as an ending time-point. E.g., "the exhibition ended on [Saturday]."
- RATE: used to classify the DATE that serves to indicate how often something happens; e.g., "[everyday]", "[every month since June]".

### 2.3.6 DEGREE

Used for academic degrees such as PhD, MBA, master etc.

**Roles**

- PERSON: E.g., "he is a newly minted [MBA]/DEGREE/PERSON".

**Sub-types** There is no sub-type for this entity type.

### 2.3.7 DISEASE

Used to classify any disease, adverse conditions/symptoms or allergy. E.g., Hodgkin's disease, AIDS, cancer, malaria, malnutrition, vomiting, skin rashes, nausea.

If the name of a virus is a cause to a disease, then mark it as DISEASE. E.g., "he was diagnosed with HIV".

**Roles** There is no role for this entity type.

**Sub-types** There is no sub-type for this entity type.

### 2.3.8 DURATION

See Section 2.2.4 for the general instructions on all temporal expressions.

Used to classify a quantified and floating amount of time, i.e., it must come with a well-known time unit, and the amount must be specified by a CARDINAL-like phrase. By floating we mean the amount of time is not anchored to any particular point in time; e.g., "3 hours" by itself is a DURATION (but "3 hours ago" is not - it's a TIME).

Anchored (non-floating) intervals should be marked as either DATE or TIME; e.g., "the [past two years]/DATE", "the [recent years]/DATE", "[from 3 to 5 in the afternoon]/TIME".

1. "Few", "several", "many, many" and the like can be used in place of a numeral.
2. Include comparative phrases, e.g., "up to two hours".
3. Sometimes after a quantifier there might be adjectives modifying a temporal unit; e.g., "9 miserable years". Include the adjective in this case.
4. It may be signaled by a preposition like "for" or "during", in which case do not include the preposition. Ex: "within [three years]", several years, a few hours, two hours, two months, two minutes, or 4 months.
5. For phrases like "2 to 5 hours" or "2 or 5 hours", mark both "2" and "5 hours" as DURATIONS.
6. "Lifetime" should not be marked since it is not quantified: no well-known time unit is used.

**Roles** There is no role for this entity type.

**Sub-types**

- RATE: used to classify the DATE that serves to indicate how often something happens; e.g., "he paid \$700 [a month]", "the tax [per year] is \$10,000".

### 2.3.9 EMAIL

Used for email address. For example: johndoe@foo.com or John Doe/Watson/IBM@IBMUS. (Lotus-styled email address)

**Roles** There is no role for this entity type.

**Sub-types** There is no sub-type for this entity type.

### 2.3.10 Events

Although an event is marked just as another mention with an appropriate event entity type in KLUE, they are governed by a set of general instructions. For this reason we shall defer their descriptions until Section 3.

### 2.3.11 FACILITY

Used for locations that are man-made structures; e.g., "a [house]", "a [restaurant]", "[Lincoln Tunnel]", "[the Golden Gate Bridge]", "[museum]", "an [apartment building]", "[pipeline]". Roads/highways/streets are FACILITY regardless if they are named or not; e.g., "[Main Street]/FACILITY". See Section 2.3.21 for comparisons between FACILITY and ORGANIZATION.

**Roles** There is no role for this entity type.

**Sub-types** There is no sub-type for this entity type.

### 2.3.12 FOOD

Used to classify anything edible and drinkable. Ex: banana cream pie, pad thai, wine.

**Roles** PRODUCT (has to be a name or a mention coref'ed with a named PRODUCT): E.g., "[Big Mac] is one of the company's signature products."

**Sub-types** There is no sub-type for this entity type.

### 2.3.13 GEOLOGICALOBJ

Used to classify all types of water bodies (Ex: sea, river, ocean, stream), mountain ranges, and other geological objects. E.g., "[Gobi Desert]", "[Rocky Mountains]", "[Mt. Everest]", "[Wake Island]".

Note if a geological object is also a geo-political entity, mark it as GPE first; otherwise mark it as GEOLOGICALOBJ. E.g., "[Taiwan]/GPE", "[Wake Island]/GEOLOGICALOBJ".

**Roles** There is no role for this entity type.

**Sub-types** There is no sub-type for this entity type.

### 2.3.14 GPE

Used to classify all types of geo-political entities such as continents ("[Europe]"), countries and their governments/people ("[American]", "[U.S.]"), government seats ("[White House]", "[Dawning 10]"), states/provinces/cities/towns and their governing bodies ("[New York]", "[City of London]", "[Westchester County]").

Note that organizations that have nations as members should not be marked as GPE; they should be marked as ORGANIZATION.MULTIGOV; e.g., United Nations, OPEC, IMF, NATO, World Bank, etc.

However, a collection of governments such as "the [west]/GPE/ORGANIZATION", "[international community]/GPE/ORGANIZATION" and "[European]/GPE.AREA/LOCATION [governments]/GPE/ORGANIZATION (relation is locatedAt) should still be marked as GPE/ORGANIZATION because they are not single organizations like NATO, OPEC etc.

When a country/location name is used to refer to the language, do not mark it. E.g., in "But for AP Spanish student Chelsea Thompson, it was a welcome challenge.", "AP Spanish" is a class for the language, so don't mark it.

#### Roles

- LOCATION: when a GPE mention is used in a geographical sense; e.g.,
  - "I visited [Spain]", "he visited [Sydney]".
  - "[Spain] was invaded".
  - "[Spain] is [where] I am from".
  - "the coast of [Britain]".

- ORGANIZATION: when a GPE mention is used to refer to the entire governing body (or bodies) of that GPE entity; e.g.,
  - "[Switzerland]'s supreme court".
  - "[US Government]".
  - "Clinton (D)": political parties is not a GPE/ORGANIZATION (it's an ORGANIZATION.POLITICAL).
  - "authorities": part of a governing body should not be marked as GPE/ORGANIZATION.
- PERSON / PEOPLE / PERSONPEOPLE: when a GPE mention is used in place of the person/people originated from that GPE entity; e.g.,
  - "most of the [Americans]".
  - "[American people]".
  - "[African]/GPE/PEOPLE [Americans]/GPE/PEOPLE" (and we can relate the two via affiliatedWith).

NOTE about the difference between role ORGANIZATION vs. role LOCATION: when the action involving the GPE in question is "organization-related", mark the mention as GPE\_ORGANIZATION. Examples:

- "[Kosovo]/GPE/ORGANIZATION announced [its]/GPE/ORGANIZATION independence from [Serbia]/GPE/ORGANIZATION".
- "[Iraq]/GPE/LOCATION was invaded".

NOTE when a GPE mention can be interpreted to serve more than one role, mark GPE (the default role). Examples:

- "pro-[Iraq] rally": can mean a rally supporting the Iraqi people or the Iraqi government.
- "[Israeli] attacks": attacks from Israeli government or people.
- "Hillary Clinton (D-[NY])": representative for that region or the people there (Clinton employedBy NY).
- "[Asian] affairs": affairs of the Asian governments or people, or simply the affairs taking place in those regions.
- "cut the supply of [Iran]'s oil": supply for the government or for the region.
- "[France] produces the best wine in the world": could be the people, or simply that region.

### Sub-types

- UNSPECIFIED: used when more than one sub-type other this one is applicable.
- AREA: used to classify a mention that refers to a geographical area that overlaps with multiple countries; e.g., "[Middle East]", "[south Asia]".
- COUNTRY: used to classify a GPE entity that is a country or an entity that has its own sovereignty; e.g., "[US]", "[Taiwan]".

### 2.3.15 LAW

Used to classify a named document that has been made into a law. E.g., "[Bill of Rights]", "[Equal Rights Amendments]", "[Treaty of Guadalupe Hidalgo]", "[Nuclear Non-Proliferation Treaty]".

1. If a nominal/pronoun mention is coref'ed with a LAW mention, mark that mention as LAW too. E.g., "the [Commodity Exchange Act] was passed in ..., the [law] mandates...".
2. If a generic term is used to refer to a specific law, but that law is not named in the article, do not mark it. E.g., "the indecency law" should not be marked if the law is not named.
3. "Shariah" should not be marked as LAW because it's not a piece of document.

**Roles** There is no role for this entity type.

**Sub-types** There is no sub-type for this entity type.

### 2.3.16 LOCATION

Used to classify all physical locations that are not already covered by another entity type (e.g., FACILITY, GPE/LOCATION or GEOLOGICALOBJ).

1. Mark direction words such as "north", "northern" etc that are used to narrow down a location or as a location: "The [north] of [Iraq]" = LOCATION GPE/LOCATION (locatedAt); "The city [north] of [Baghdad]" = LOCATION GPE/LOCATION (locatedAt); "[northern] [Iraq]" = LOCATION GPE/LOCATION (locatedAt).
2. "Boarder" is a LOCATION.

**Roles** There is no role for this entity type.

**Sub-types** There is no sub-type for this entity type.

### 2.3.17 MEASURE

Used to classify measurement units: distance, energy, speed, temperature, volume, and weight. Amount must be included or implied. E.g., "[4 miles]", "[4 grams]", "[4 degrees]", "[4 pounds]", "[4 ounces]", "\$4 [per gallon]" ("4" is MONEY and "per gallon" is a MEASURE), "[per barrel]".

1. Just like temporal expressions (Section 2.2.4), always try to find the maximal extent for a MEASURE mention. E.g., "[15 to 18 miles]".
2. Sometimes a number/comparative phrase is left out, but it is implied in the context. E.g., "he traveled [miles] to get here" - we still mark "miles" as MEASURE.
3. Treat these as whole measures: "[60 miles per hour (mph)]", "[30 miles per gallon (mpg)]", "[2.3 parts per million (ppm)]".

**Roles** There is no role for this entity type.

**Sub-types** There is no sub-type for this entity type.

### 2.3.18 MONEY

Used to classify values including all monetary denominations. The denomination (word or symbol) should be included in the mention if present. E.g., "[50 yen]", "[one million dollars]", "[17,000 pounds]", "\$10.20".

1. Just like temporal expressions (Section 2.2.4), always try to find the maximal extent for a MEASURE mention. E.g., "[from 75 to 100 dollars]".
2. A numeric amount must be included or implied; e.g., "the dollar lost value against the yen" - neither dollar or yen should be marked.

**Roles** AWARD: E.g., "the prize is [\$3 million]".

## Sub-types

- UNSPECIFIED: used when more than one sub-type other than this one is applicable.
- WORTH: used to classify money expressions describing the worth of a person/organization; e.g. "Bill Gates's total net worth is [\$40 billion]/MONEY.WORTH for 2009.". DO NOT INCLUDE stock price.
- GAINLOSS: used to classify money expressions describing gain/loss, including revenue, profit, loss etc of a person/organization; e.g. "Despite losing [\$18 billion]/MONEY.GAINLOSS in the past year, Gates is still the richest person in the world". DO NOT INCLUDE stock price changes.

### 2.3.19 ORDINAL

Used to classify all ordinal numbers if they refer to the ordering of things in the context (e.g., "first", "second" etc).

1. Mark the ordinal word separately from the word it modifies, e.g. "the [first] [round] of the [tournament]".
2. In "No. 12", "12" is an ordinal, but not "No."
3. Mark "last" and "final" as ORDINAL when they are used to specify ordering (don't mark "final" in "final decision").
4. If the ordinal word stands by itself and it refers to something else, mark it with ORDINAL but assign an appropriate role E.g., "Jeneba Jalloh Ghatt is the [first]/ORDINAL/PERSON to admit that she likes nice things".

**Roles** Every entity type, except for AGE / CARDINAL / DATE / DURATION / ORDINAL / PERCENT / TIME, can be a role for ORDINAL. E.g., "the [first]/ORDINAL/PERSON of the contestants", "the [second]/ORDINAL/VEHICLE of the cars".

**Sub-types** There is no sub-type for this entity type.

### 2.3.20 ORGAN

Used to classify an organ or a body part (e.g., head) of a human or animal. In biology, an organ is a group of tissues which perform some function. Common human and animal organs include the heart, lungs, brain, eye, stomach, spleen, pancreas, kidneys, liver, intestines, skin, uterus, bladder, bone, etc. If an ORGAN is used in a figurative way, such as "he is the brain of the group", do not mark it as ORGAN. Do not mark a person's body/remains as ORGAN. See PERSON (Section 2.3.24).

**Roles** There is no role for this entity type.

**Sub-types** There is no sub-type for this entity type.

### 2.3.21 ORGANIZATION

Used to classify hierarchical social structures (including companies) that are not GPE, including the organizations whose members are countries; e.g. "[IBM]", "the [Yankees]", "the [Girl Scouts]", "the [New York Times]", "[NATO]", "[OPEC]".

1. Comparison to FACILITY (Section 2.3.11): a FACILITY is simply a man-made structure – it cannot be an agent of an event (e.g., "announced", "said"), and it cannot be an recipient of a non-physical action such as "criticism".
2. Mark factions, groups, movements if the context indicates they should be interpreted as ORGANIZATION.
3. Mark "Wall Street", "Hollywood" as ORGANIZATION because they refer to the entire industry.

4. Bodies such as Islam should also be marked as ORGANIZATION.RELIGIOUS.
5. A fund's name should be marked as ORGANIZATION if it is the name of the organization, e.g. "Long-term Capital".
6. In sports articles, always mark "league" as ORGANIZATION, even though sometimes it is used to refer to a set of games.

### Roles

- FACILITY: E.g., "He was transferred to [St. Joseph's Hospital]", "I went to [IBM] today", "People gathering at [IRIB] (Islamic Republic of Iran Broadcasting) to protest against its biased broadcast."
- PEOPLE: E.g., "[Police] arrested the suspect."
- PRODUCT: E.g., "they like to use the search engine [Google]"
- TICKER: E.g., "[MSFT] announced the intent to buy FAST for \$1.2 billion."
- TITLWORK: E.g., "the latest issue of [New York Times]..."

### Sub-types

- UNSPECIFIED: used when more than one sub-type other this one is applicable.
- MULTIGOV: multi-government organizations (members are governments); e.g., "[UN]", "[OPEC]", "[NATO]", "[IMF]", "[World Bank]".
- GOVERNMENT: government organizations; e.g., "[DOD]"
- COMMERCIAL: for-profit organizations; e.g., "[IBM]" etc.
- EDUCATIONAL: educational organizations; e.g., "[Yale]" etc.
- POLITICAL: political organizations; e.g., "[Democratic Party]" etc.
- RELIGIOUS: religious organizations; e.g., "[Church of Scientology]"
- SPORTS: sports-related organizations; e.g., "[sports teams]", "[NBA]", etc.
- MILITARY: all armed forces, including militia; e.g., "[Army]"

### 2.3.22 PEOPLE

1. Used to classify a group of people except people from a geo-political entity (see GPE for comparisons). E.g., "[Christians]", "[Islamists]"
2. Used in the case of a family name when speaking of more than one member of the family. E.g., "[The Petersons]"
3. Used for troops or forces, sources, delegations, factions, committees when there is doubt to their organizational integrity.
4. Words describing people with certain ideological bent can be marked as PERSON or PEOPLE depending on what entity they refer to. Examples:
  - "[Islamist] organization", "[Jewish] organizations", "[Maoist] organization": PEOPLE (relation is affiliatedWith).
  - "he is a [separatist]", "he is a [Maoist]": PERSON (coref).
  - "he is [Maoist]", "he is [Jewish]": PEOPLE (relation is partOfMany and affiliatedWith).
  - "he is a [republican]", "he is a [Catholic]": PERSON (coref).



## Roles

- OCCUPATION: when a mention refers to a job title (see the link for more info).

**Sub-types** There is no sub-type for this entity type.

### 2.3.23 PERCENT

Used to classify percentage. Percent symbol or the actual word percent must be marked also. Ex: 50%, three quarters of a percent, seventy percent, 'a few percentage points'.

1. Words like "few", "many" or "several" are also included; e.g. "a few percentage points" should be tagged as one mention PERCENT.
2. For comparatives like "[more than 20%]", include the comparative phrase.
3. For percent expressions referring to another type of entities, mark it as PERCENT but choose an appropriate role. E.g., "[20%]/PERCENT/PEOPLE of the people failed passing the test".

**Roles** Every entity type, except for AGE / CARDINAL / DATE / DURATION / ORDINAL / PERCENT / TIME, can be a role for PERCENT. E.g., "[20%]/PERCENT/PEOPLE of the population".

**Sub-types** There is no sub-type for this entity type.

### 2.3.24 PERSON

Used to classify an individual, fictional or real. The mention can be the person's name, whether preceded by a title or not ("[John Smith]", "[Dr.] [Smith]"), a nominal, non-occupational reference to the person ("[winner] of the Nobel Prize"), or an occupational, age, salutation reference to a person ("A [lawyer] was walking down the street", "the [30-year-old] won the prize").

1. For PERSON mentions that are titles with a GPE mention, make sure to separate the GPE from occupation. E.g., "John Smith, the [US] [special representative] in Iraq" or "George Bush, the [president] of the [United States]", "the [NY] [attorney general]".
2. A person's body/remains should also be marked as PERSON and coref'ed to his/her name if it appears (but specific organs should still be marked as ORGAN).

## Roles

- OCCUPATION: when a mention refers to a job title (see the link for more info).

**Sub-types** There is no sub-type for this entity type.

### 2.3.25 PERSONPEOPLE

Similar to PERSON and PEOPLE. Used when the plurality is unknown (this occurs in languages such as Chinese, which lacks morphological indication of number).

## Roles

- OCCUPATION: when a mention refers to a job title (see the link for more info).

**Sub-types** There is no sub-type for this entity type.

### 2.3.26 PHONE

Used for phone numbers; e.g., (914)123-4567, 1-800-123-4567, 1-800-ASK-HELP.

**Roles** There is no role for this entity type.

### **Sub-types**

- FAX: Used for FAX numbers.

### **2.3.27 PLANT**

Used to classify any plant or plant-like object that has a name. Ex: Black eyed Susan, Peonies, and Joshua Tree.

### **Roles**

- SUBSTANCE: E.g., "Two arrested people in Kazakhstan, said that they gathered 603 kg of [marijuana] for personal use."

**Sub-types** There is no sub-type for this entity type.

### **2.3.28 PRODUCT**

Used to classify any profitable named thing that can be a trade name, brand name or trademark, and which in the context does not fall under any other entity type (i.e., PRODUCT is the fallback type for trade names/brand names/trademarks). Examples:

1. "Windows", "TOEFL", "iTunes", and "Cabbage Patch Doll" are all PRODUCTS because they are trademarks/brand names and there is no other entity type that can be applied to them.
2. "Viagra" should be marked as SUBSTANCE (optionally with role PRODUCT) because it is a substance first. "[Toyota] [Land Cruiser]" should be marked as ORGANIZATION VEHICLE instead of ORGANIZATION PRODUCT for a similar reason.
3. Just like LAW, if a nominal/pronoun mention is coref'ed with a PRODUCT mention, mark that mention as PRODUCT too. E.g., "When [Windows] first came out, [it] didn't have true multitasking capability."
4. In "the search engine [Google]", we mark it as ORGANIZATION.PRODUCT because "Google" in most contexts is a company. When combined with a company's name, make sure to separate the product name from the company name. E.g., "[Microsoft] [Windows]" = ORGANIZATION PRODUCT.

**Roles** There is no role for this entity type.

**Sub-types** There is no sub-type for this entity type.

### **2.3.29 SUBSTANCE**

Used to classify drugs, chemicals and natural substance such as elements. It must be significant in the document. E.g., "[boron]", "[AZT]", "[Viagra]", "[oil]", "[natural gas]", "[oxygen]", "[gold]", "[aluminum]", etc.

### **Roles**

- PRODUCT: "Roche may boost production of [Tamiflu]..."
- WEAPON: "In 2001 [anthrax] was deployed as a biological weapon via mail using the U.S. Postal system."

**Sub-types** There is no sub-type for this entity type.

### 2.3.30 TICKER

Used for stock tickers ONLY; e.g., "MSFT" for Microsoft, "JAVA" for Sun, "IBM" for IBM (note for the last two examples you need to be sure that they refer to the stock ticker symbols, not other things such as the programming language and the company).

**Roles** There is no role for this entity type.

**Sub-types** There is no sub-type for this entity type.

### 2.3.31 TIME

See Section 2.2.4 for the general instructions on all temporal expressions.

Used to classify a time expression; e.g., "[Monday morning]", "[3pm]", "[10 o'clock].

1. The a.m. and p.m. should be included in a TIME mention if exists.
2. See Section 2.2.4 for event-dependent time expressions.

**Roles** There is no role for this entity type.

#### Sub-types

- **STARTING**: used to classify the TIME that serves as a starting time-point. E.g., "the game started at [6pm]."
- **ENDING**: used to classify the TIME that serves as an ending time-point. E.g., "the exhibition ended at [around 10pm]."
- **RATE**: used to classify the TIME that serves to indicate how often something happens; e.g., "[every hour]".

### 2.3.32 TITLWORK

To classify named titles of creative works like books, songs, and television shows. They may be surrounded by quotation marks in the text, but DO NOT MARK THE QUOTATION MARKS. E.g., "[Empire Strikes Back]", "[Moby Dick]", "The [Bible]", "[Blue Moon]".

#### Roles

- **PRODUCT**: "the [Harry Potter series] has sold more than 400 million copies to date"

#### Sub-types

- **PUBLICATION**: "[Bible]"

### 2.3.33 VEHICLE

Used to classify physical devices that transport an object from one location to another. It includes all air, land, and water vehicles, such as airplanes, cars, bicycles, and trains.

#### Roles

- **PRODUCT**: "The 2009 Honda [CR-V] is still America's most-popular compact SUV"
- **WEAPON**: "the [plane] crashed into the buildings..."

**Sub-types** There is no sub-type for this entity type.

### 2.3.34 WEAPON

Used to classify devices that can physically harm or destroy people or facilities. It includes blunt weapons ("a [baseball bat]" if it's used as a weapon), explosives, sharp weapons ("[knife]"), chemical weapons, biological weapons, nuclear weapons, and shooting devices ("[gun]"). Also mark weapon names such as "[Ak-47]", "[Al-Samoud]", "[AIM-54 Phoenix]".

**Roles** There is no role for this entity type.

**Sub-types** There is no sub-type for this entity type.

### 2.3.35 WEATHER

Used to classify specific non-disastrous weather conditions, such as El Nino, global warming, and the greenhouse effect.

**Roles** There is no role for this entity type.

**Sub-types** There is no sub-type for this entity type.

### 2.3.36 WEB

Used for website addresses. A web address does not need to include the `http://` or `https://` prefix, but it has to be an address. E.g., `http://www.ibm.com` and `ibm.com` are both web addresses, but `IBM` is not.

**Roles**

- ORGANIZATION: "[Facebook.com] announced their membership reached 250 millions."

**Sub-types** There is no sub-type for this entity type.

## 3 Events

An event is something that happens/happened/will happen/might happen in a specific time. *Note we do not mark a generic event*; i.e., events that cannot be pinpointed to a specific time should not be marked. Examples:

- "Hurricanes usually come in summers": "Hurricanes" should not be marked since "summers" are not specific.
- "A [hurricane]/EVENT\_DISASTER came last summer": "hurricane" is marked since we have a specific time.

Linguistically an event can be realized by a verb (e.g., "killed"), a gerund (e.g., "killing"), a noun (e.g., "discussion"), or a phrase with mixed grammatical categories (e.g., "fling suit").

Hypothetical events should be annotated; E.g., "he might participate in that [meeting]/EVENT\_MEETING". However, we do not mark negative events; E.g., "he did not attend that meeting" should not be marked.

An event can be small-scale (e.g. a fist fight) or large-scale (a war). Don't mark small-scale sub-events, like a strikeout in a ball game, or a punch in a fight.

### 3.1 General Instructions

#### 3.1.1 Phrasal Verbs

A phrasal verb is a phrase (more than one word) that is used as a single verb. E.g., "he [called for]/EVENT\_COMMUNICATION an immediate cease-fire". All of the words in a phrasal verb should be marked if the phrasal verb is an event.

In case of a discontinuous event phrase, e.g. "locked them up", mark only the first verb "locked". This happens less frequently in English.

### 3.1.2 Roles

Only a few events can act as another type of entity and therefore they have roles. These are listed in Section 3.2 below.

### 3.1.3 Relations involving Events

All (except one) events can be arguments to the following KLUE relations (See Section 5 for the relation guidelines):

- **affectedBy**: whom did it affect?
- **agentOf**: who initiated the event? For `EVENT_AWARD` use the relation `awardedBy` for its agent.
- **before**: what was the time/event an event happened before?
- **locatedAt**: where did it happen?
- **near**: near where did it happen?
- **overlaps**: what was another event that this event overlaps with in time?
- **participateIn**: who was involved?
- **partOf**: was it part of a larger-scale event, e.g. Battle of Blil Run and Civil War?
- **partOfMany**: was it one of several events, e.g. WW1 and 20th century wars?
- **timeOf**: when did it happen?

Some events can also be arguments to relations that are not listed above; these events and the relations they can involve in are given in the section Definition of Event Types below.

## 3.2 Definition of Event Types

### 3.2.1 `EVENT_AWARD`

These are the events where a prize/academic degree was/will be awarded.

Examples:

- "In 2007, [Gore]/PERSON was [awarded]/EVENT\_AWARD the [Nobel Peace Prize]/AWARD":
  - "Gore" affectedBy "awarded";
  - "Nobel peace prize" awardedTo "Gore".
- "After [graduating]/EVENT\_AWARD with a [Juris Doctor]/DEGREE ([J.D.]/DEGREE) from [Harvard]/ORGANIZATION.EDUCATIONAL in 1991, [he]/PERSON returned to Chicago":
  - "he" affectedBy "graduating";
  - "Harvard" agentOf "graduating";
  - "Juris Doctor" / "(J.D.)" awardedTo "he";
  - "Juris Doctor" / "(J.D.)" awardedBy "Harvard".
- "[Lehman]/ORGANIZATION.COMMERCIAL was [awarded]/EVENT\_AWARD Best Investment Bank by [Euromoney]/ORGANIZATION":
  - "Lehman" affectedBy "awarded";
  - "Euromoney" agentOf "awarded";
  - "Best Investment Bank" awardedTo "Lehman";
  - "Best Investment Bank" awardedBy "Euromoney".

### 3.2.2 EVENT \_ BUSINESS

The normal activities (mostly financial) of an organization: making or losing money, changing ownership, expanding, or lending or borrowing money.

### 3.2.3 EVENT \_ COMMUNICATION

1. An intentional act of imparting the agent's ideas via words (can be verbal or in writing), e.g. telling, saying, talking, reporting, emphasized, pointed out, argued, threatened, expressed hope, expressed concern, citing, quoting, publishing.
2. If there is doubt whether the information-imparting act is intentional or can only be done via words, do not mark it.
3. Do not mark "book" because it is not necessarily meant to be communicating the author's ideas (a book can be summarizing someone else's ideas or telling a story etc).
4. When a verb indicating a communication event co-occurs with a noun that also indicates the same event, mark only the verb; e.g., "he [said]/EVENT \_ COMMUNICATION in the report".
5. If the only words indicating a communication event is a noun, mark the noun as EVENT \_ COMMUNICATION; e.g., "according to his [statement]/EVENT \_ COMMUNICATION".

More examples:

- "based on his [account]/EVENT \_ COMMUNICATION".
- "[Responding]/EVENT \_ COMMUNICATION to a [statement]/EVENT \_ COMMUNICATION...": the "statement" is marked because it is a different event from "responding".
- "Intel [made]/EVENT \_ COMMUNICATION a statement in response to...": do not mark "statement" because we have the verb "made".
- "He [broke]/EVENT \_ COMMUNICATION the news...": do not mark "news" because we have the verb "broke".
- In "He [said]/EVENT \_ COMMUNICATION in a letter [quoting]/EVENT \_ COMMUNICATION Ben...", coref "said" and "quoting" (they are same events), but in "he is [quoted]/EVENT \_ COMMUNICATION as [saying]/EVENT \_ COMMUNICATION" do not coref the two because they are two different events (their agents are different).
- Do not mark singing, chanting, etc.

### 3.2.4 EVENT \_ CRIME

Used to mark an action or an instance of negligence that is deemed injurious to the public welfare or morals or to the interests of the state and that is legally prohibited. Note that:

1. Crime includes not only violent crimes but also non-violent crimes such as embezzlement, corruption, etc.
2. As long as the context indicates it is a crime it should be marked as such, regardless of your preconception about what constitutes as crime. E.g., homosexuality is a crime in some culture/country - if the context indicates so mark it as a crime.
3. For something to be a crime the context must indicate/imply that it is legally prohibited. E.g., if police is after a person because he was the agent of an event, that event most likely is a crime.

**EVENT\_CRIME vs. Other Events** When judging from the sentential context of an event, if the event can be marked as an `EVENT_CRIME` or one of the other event types, `EVENT_CRIME` trumps. E.g.,

- "An 18-year-old California man has been charged with [kidnaping]/`EVENT_CRIME` a 15-year-old Idaho girl..." (if there is no indication this is a crime, "kidnaping" should be marked as `EVENT_CUSTODY`).
- "he has been accused of [beating]/`EVENT_CRIME` his dog" (if there is no indication this is a crime, "beating" should be marked as `EVENT_VIOLENCE`).

### 3.2.5 `EVENT_CUSTODY`

An arrest, imprisonment, release, escape, etc... This can be carried out by either the authority, such as police or by other non-authoritative organization. For instance, "the insurgents kidnaped three Japanese", or "the insurgents released the three Japanese hostages" should be both annotated as `EVENT_CUSTODY`.

See Section 3.2.4 to decide if an event is an `EVENT_CRIME` instead.

### 3.2.6 `EVENT_DEMONSTRATION`

A rally, parade, or protest march. The category also includes strikes and walkouts against an organization.

### 3.2.7 `EVENT_DISASTER`

Events that caused damage, such as an accident (e.g. train crash), natural disaster (e.g. earthquake), or epidemic (e.g. Ebola outbreak). If the text does not explicitly mention damage, mark when it's clear that it usually causes damage.

1. Mark only the event itself and not its consequences (e.g. killing).
2. `EVENT_DISASTER` vs. `WEATHER`: mark `EVENT_DISASTER` if damage is explicitly mentioned (e.g., "heavy [rain]/`EVENT_DISASTER` caused houses to [collapse]/`EVENT_DISASTER`").

### 3.2.8 `EVENT_EDUCATION`

These are the events where a person was/will receive education/training. Note this includes all kinds of training, not just schooling.

Examples:

- "After [graduating]/`EVENT_EDUCATION` with a Juris Doctor (J.D.) from [Harvard]/`ORGANIZATION.EDUCATION` in 1991, [he]/`PERSON` returned to Chicago": in addition to the `EVENT_EDUCATION`, "he" is the agentOf the event, and "he" is related to "Harvard" via the relation `educatedAt`.
- "[he]/`PERSON` completed his [college education]/`EVENT_EDUCATION` in [England]/`GPE.LOCATION`": here the `EVENT_EDUCATION` is "college education", "he" is the agentOf the event, and the event is related to "England" via the relation `locatedAt`.

In addition to the relations listed in Section 3.1.3, this event can be also an argument to the following relation:

- `educatedAt`: in what organization did/will be the education take place?

### 3.2.9 `EVENT_ELECTION`

Used to mark events related to elections (including elections for non-governmental positions such as chairman of a board), including media coverage, inauguration, voter turnouts, election results, announcements of nominations (for candidacy) and of a candidate's intention, polls, projections, confirmation (if the context indicates it's the result of an election), and campaign-related events.

### 3.2.10 EVENT\_LEGAL

Any event that is a result of the involved person/organization's illegal conduct should be marked as EVENT\_LEGAL. Examples:

- filing (a lawsuit), prosecution, charge, allegation
- sentencing, execution, indictment, fine
- extradition
- court case, court order, lawsuit, trial
- sanctions (against a country); e.g., "UN Liberia [sanctions]/EVENT\_LEGAL, [adopted]/EVENT\_LEGAL by a Security Council resolution in March, are to go into effect today..."
  - "sanctions" and "adopted" are two events: the former is like "verdict", the latter is like "sentencing".
  - Following the reasoning above both should be EVENT\_LEGAL because they are result of illegal conduct of a specific person/organization (in this case a government).

Note:

1. Examples of EVENT\_LEGAL in both verb and noun
  - "[filed]/EVENT\_LEGAL a [sexual harassment]/EVENT\_CRIME [lawsuit]/EVENT\_LEGAL" and then coref the two EVENT\_LEGALs.
  - "[filed suit]" because both are the same event.
2. Legal pronouncements (e.g. a court order) should be EVENT\_LEGAL rather than EVENT\_COMMUNICATION.
3. Execution (for crimes) should be EVENT\_LEGAL rather than EVENT\_VIOLENCE.
4. Do not mark crimes; use EVENT\_CRIME instead.
5. Do not mark the creation/passing/signing of laws/public policies; use EVENT\_LEGISLATION instead.
6. Do not mark marriage and divorce.
7. Implication to relations: when someone is extradited from Country X to Country Y, the event is locatedAt X, and the person was locatedAt Y in the past.

### 3.2.11 EVENT\_LEGISLATION

Used to mark creation/passing/signing of laws/bills/public policies. Examples:

- "The group also hopes to persuade legislators to [pass]/EVENT\_LEGISLATION laws banning the behavior."
- "A national safety group is seeking a nationwide ban on cell phone use while behind the wheel." - do not mark "seeking" or "ban", as it is not an event of a bill/law passing.
- "It is unclear if the democrat's bill will [pass] if the republicans reject it" - do not mark "reject" because it is negative.

### 3.2.12 EVENT\_MEETING

Used to mark events where multiple people participate in real-time to exchange information among themselves verbally. E.g., "[Parliamentary session]", "[general assembly]", "FBI [interrogation]", "the agent is interrogating the man".



**EVENT\_MEETING vs. EVENT\_COMMUNICATION** if the act of "imparting information" is bidirectional, it is a meeting. E.g., "debate" is an **EVENT\_MEETING**.

### 3.2.13 EVENT\_PERFORMANCE

A scripted ceremony or performance that must have a live audience (e.g. graduation, religious service, dedication, theatrical performance).

1. Compared with **TITLEWORK**: if in the context there is an indication that there is a live audience then mark the mention as **EVENT\_PERFORMANCE**, otherwise consider **TITLEWORK**. E.g., if the context indicates/implies there is a live audience for "American Idol", it should be marked as **EVENT\_PERFORMANCE**, even we know it's recorded as a TV show as well.
2. Funerals and weddings are **EVENT\_PERFORMANCE**.
3. A movie is not an **EVENT\_PERFORMANCE** because it doesn't have a live audience.
4. "John played in [Woodstock]/**EVENT\_PERFORMANCE**": John is participantIn Woodstock.

### 3.2.14 EVENT\_PERSONNEL

A possible change in employment (including changes of membership of a sports team), e.g. hiring, firing, resigning and nominating (except for elections - use **EVENT\_ELECTION** instead).

In addition to the relations listed in Section 3.1.3, this event can be also an argument to the following relation:

- **employedBy**: in a hiring event, this relation relates a person to the organization s/he is/was hired by.

### 3.2.15 EVENT\_SPORTS

A game or set of games (as in a tournament). A sports event must be independently winnable, e.g. a leg of a bicycle race but not an inning in baseball or a hole in golf.

#### Roles

- **AWARD**: mark with this role when a sports event is used to refer to the award/prize; e.g., "they won the [World Cup]/**EVENT\_SPORTS/AWARD**".

### 3.2.16 EVENT\_VIOLENCE

An act of physical or military hostility, e.g. shooting, spraying with Mace, battle, Second World War. If there is doubt about the hostility, do not mark it. E.g., in "the FBI raided the office.." don't mark "raid" (mark **EVENT\_LEGAL** in this context). One test is to ask if the event participants can be called enemies among themselves.

Do not mark the product of the event if it co-occurs with the event; e.g., in "they were killed in the war" don't mark "killed". But do mark the product of the event if the event is not mentioned, e.g., in "they were killed during the time", if from the context it is clear it's a result of an **EVENT\_VIOLENCE** and the event itself is not mentioned, mark "killed" as **EVENT\_VIOLENCE**.

See Section 3.2.4 to decide if an event is an **EVENT\_CRIME** instead.

In addition to the relations listed in Section 3.1.3, this event can be also an argument to the following relation:

- **instrumentOf**: the instrument by which the event was carried out.

### 3.2.17 EVENT

It is used to mark an event argument participating in a temporal relation (before or overlaps) if the argument cannot be marked with another more specific event type (see Section 2.2.4 on marking event-dependent temporal expressions). In addition to the relations listed in Section 3.1.3, this event can be also an argument to the following relation:

- **instrumentOf**: the instrument by which the event was carried out.

## 4 Entities (Coreference)

The goal of coreference annotation is to identify mentions that are equivalent, i.e., they refer to the same entity. Such mentions are linked together into a so-called “coreference chain”. For instance, “Mary is beautiful and everybody likes her”, “Mary” and “her” in this sentence refer to the same entity (i.e., Mary) and should be linked into the same coreference chain.

### 4.1 General Instructions

1. We coref mentions regardless of their entity types, as long as it makes sense to a human reader. For example, "one/CARDINAL of them was the author/PERSON of the book".

NOTE this is different from the previous incarnation of this guidelines - we used to allow only limited cross-type coreference.

2. For mentions belonging to the following entity types, we do not annotate coreference relationships:
  - EMAIL.
  - MEASURE.
  - PHONE.
3. Pay attention to actor/actress and the role s/he plays: never coref the actor/actress with the role - use playsRoleOf instead.

## 5 Relations

Two entities are in a relation if they meet the semantic criteria for the relation (see below) AND there is direct syntactic evidence (not just world knowledge) for the relation. Some evidence types are prepositions ("officials in California"), possessive markers ("John's mother"), apposition ("Jane Clayson, ABC News"), and arguments of the same verb ("Adam Smith visited Japan").

We mark a relation even if it is hypothetical (see Section 5.1.6 on tense/modality/specificity class for instructions). However we do not mark negative relations; e.g., in "[he]/E1 did not [visit]/E2 Paris" the relation E1 agentOf E2 should be discarded.

For each relation marked we also need to mark its tense, modality and specificity class and optionally its anchor words (see Section 5.1.6 and Section 5.1.7).

### 5.1 General Instructions

#### 5.1.1 Argument Type Restrictions

Every relation accepts two arguments (arg1 and arg2) of compatible types. For example, arg2 of relation bornOn must be of a time-related type - other types such as VEHICLE is not allowed (and clearly not making sense).

When considering compliance of argument type restrictions, if roles are present for the arguments, the roles can be used to satisfy the restrictions. For example, while CARDINAL parentOf PERSON is not allowed ("one of them is his father"), if arg1 is annotated as having role PERSON, then the arguments satisfy the argument type restriction of relation parentOf.

These compatible types are indicated in Section 5.2.

#### 5.1.2 Multiple Choices of Mention Pairs and "Most Logical" Rule

If you have to choose among multiple mention pairs to mark a relation, pick the pair that are most logical. E.g.,

- in "[John]/PERSON has been working for [IBM]/ORGANIZATION in [his]/PERSON position for many years" we mark "John" employedBy "IBM" instead of "his" employedBy "IBM";

- in "[John Doe]/PERSON, although [he]/PERSON is in [jail]/FACILITY, remains married to [Mary Smith]/PERSON" we mark "John Doe" spouseOf "Mary Smith" instead of "he" spouseOf "Mary Smith" and "he" locatedAt "jail";
- in "[Marlins]/ORGANIZATION [pitcher]/PERSON [Hernandez]/PERSON" we mark "Hernandez" memberOf "Marlins" instead of "pitcher" memberOf "Merlins".

### 5.1.3 Symmetric Relations and Transitive Relations

Some relations are symmetric (e.g. two sisters are relatives of each other) or transitive (e.g. if Albany is locatedAt New York, and New York is locatedAt the United States, then Albany is locatedAt the United States). These are indicated in the section Definition of Relation Types below.

### 5.1.4 Implied Relations

Some relations can be implied by some other relations: the implied relations can come from a symmetric relation (John relative Mary implies Mary relative John), a chain of transitive relations (Albany locatedAt New York and New York locatedAt US, thus Albany locatedAt US), or by the semantics of the relations – some possibilities are listed below (they are by no means comprehensive – see examples below):

1. **affectedBy, citizenOf, clientOf, colleague, educatedAt, employedBy, founderOf, managerOf, memberOf, ownerOf, parentOf, partner, partOf, relative, shareholdersOf, siblingOf, spokespersonFor, spouseOf** and **subsidiaryOf**: all imply affiliatedWith.
2. **agentOf**: implies affectedBy.
3. **basedIn**: implies locatedAt.
4. **bornAt**: implies residesIn and locatedAt (at one point in time).
5. **capitalOf**: implies locatedAt.
6. **diedAt**: implies locatedAt (at one point in time).
7. **diedOf**
  - When the 2nd argument is DISEASE it implies hasAttribute.
  - When the 2nd argument is EVENT\_CRIME, EVENT\_DISASTER, EVENT\_VIOLENCE or EVENT, it implies affectedBy.
8. **educatedAt**: implies memberOf and locatedAt.
9. **employedBy**: implies memberOf when the 2nd argument is ORGANIZATION, GPE or WEB.
10. **founderOf**: implies memberOf.
11. **managerOf**: implies memberOf.
12. **parentOf**: implies relative.
13. **residesIn**: implies locatedAt.
14. **siblingOf**: implies relative.
15. **spokespersonFor**: implies memberOf ("employedBy" is not listed here because the person might get no monetary reward from the organization).
16. **spouseOf**: implies relative.

As a principal, we only mark the relations that are explicitly specified in the text, and when multiple relations apply, we always try to pick the most specific one to mark, i.e., the one that can imply the other candidates. Examples:

- For a symmetric relation such as colleague, do not create two relations between the same pair of mentions; only create one relation based on the order of occurrence of the two mentions; e.g., for "Mary is married to John" only create relation Mary RELATIVE John, not vice versa.
- For a chain of transitive relations, relate the adjacent entities in the chain and skip the ones that relate the non-adjacent entities; e.g, "Albany, NY, US" we mark only Albany locatedAt New York and New York locatedAt US, but not Albany locatedAt US.
- For a mention pair that multiple relations apply, pick the ones that are most specific to mark; e.g., "[Jobs]/E1 is the CEO of [Apple]/E2" mark E1 managerOf E2 and skip E1 memberOf E2 (more examples are given here).
- From the sentence "the band features guitarist Ray Jackendoff, drummer Peter Culicover..." we know both players are memberOf the band. But by inference we also know they are colleague of each other. But we should only mark memberOf, because it is more readily conveyed from the text.

### 5.1.5 Multiple Relations Between the Same Mention Pair

In general more than one relations can hold between a pair of entities, but no more than one relation of the same type can hold between a pair of entities.

If more than one relation apply between the same pair of arguments:

- If one of the candidates can be determined to be more specific than the others, i.e., it can imply that the other relations hold, choose the most specific relation. See here for a list of relations and their implied relations. Examples:
  - In "John lives in Yorktown" both locatedAt and residesIn apply between "John" and "Yorktown", but we choose residesIn because it implies locatedAt (thus it is more specific).
  - In "Salim's research group" both memberOf and managerOf apply between "Salim" and "research group", but we choose managerOf because it is more specific.
  - In "John, the spokesperson of the company", relation "memberOf" and "spokespersonFor" all apply, but we choose spokespersonFor because it implies all the others.
- If none of the relations imply the others, mark all of them. E.g., "[He]/E1 passed away in his [birthplace]/E2": both E1 diedAt E2 and E1 bornAt E2 hold.

### 5.1.6 Tense, Modality and Specificity Class

For every relation annotated we also need to mark its tense, modality and specificity class:

**Tense** This marks the time the relation holds with respect to the publication time of the document. There are four possible values: PAST, PRESENT, FUTURE and UNSPECIFIED.

1. PAST: the relation is taken to hold only for some span prior to the publication time.
2. PRESENT: the relation is taken to hold for a limited time overlapping with the publication time.
3. FUTURE: the relation is taken to hold only for some span after the publication time.
4. UNSPECIFIED: the relation is either static or the span of time for which it holds cannot be determined with certainty. Note that the following relations should always be marked as UNSPECIFIED for tense (they are also marked in Section 5.2 below):

- affectedBy
- agentOf
- authorOf
- before
- bornAt

- bornOn
- diedAt
- diedOf
- diedOn
- dissolvedOn
- educatedAt
- founderOf
- foundedOn
- instrumentOf
- overlaps
- parentOf
- participantIn
- partOf only when both of the arguments are events.
- partOfMany
- siblingOf
- timeOf

Examples:

- "[Wolf Blitzer]/E1, [CNN]/E2": E1 employedBy E2 is PRESENT.
- "[they]/E1 [told]/E2 reporters": E1 agentOf E2 is PAST.
- "the product was manufactured in the [Peanut Corp. of America]/E1 [plant]/E2": E1 ownerOf E2 is PRESENT.
- "the [president-elect]/E1 of the [US]/E2": E1 managerOf E2 is FUTURE.
- "[my]/E1 [daughter]/E2": E1 relative E2 is UNSPECIFIED.
- "[Yorktown]/E1, [NY]/E2": E1 locatedAt E2 is UNSPECIFIED.
- "[Los Angeles]/E1 [Police Department]/E2: E2 partOf E1 is UNSPECIFIED.
- "[Microsoft]/E1, the [software maker]/E2": E1 hasAttribute E2 is PRESENT (because company role may change over time).

**Modality** This marks whether the relation was asserted in the document as a fact, rather than a hypothesis/speculation. There are two possible values: ASSERTED and OTHER.

1. ASSERTED: when the relation can only be interpreted to hold in the world the arguments exist in (see the two examples about movie characters below).
2. OTHER: when the relation can be interpreted to hold in a counterfactual world - a world that the arguments do not exist in.

Examples:

- "the product was manufactured in the [Peanut Corp. of America]/E1 [plant]/E2": E1 ownerOf E2 is ASSERTED.
- "[Wolf Blitzer]/E1, [CNN]/E2": E1 employedBy E2 is ASSERTED.
- "[US]/E1 [president candidate]/E2": E2 managerOf E1 is OTHER.
- "the [president-elect]/E1 of the [US]/E2": E1 managerOf E2 is ASSERTED.

- "the salmonella could have come from a [sick worker]/E1 [there]/E2": E1 locatedAt E2 is OTHER.
- "We are afraid that Al-Qaeda [terrorists]/E1 will be in [Baghdad]/E2": E1 locatedAt E2 is OTHER.
- "If they can get tickets today, [they]/E1 will be in [Baghdad]/E2 tomorrow": E1 locatedAt E2 is OTHER.
- "In the movie Marlon Brando was saved by [his]/E1 youngest [son]/E2": E1 relative E2 is OTHER (it is the character in the movie whose son saved him).
- "In the movie [Don Corleone] was saved by [his]/E1 youngest [son]/E2": E1 relative E2 is ASSERTED.

**Specificity Class** This marks whether a relation is specific. There are two possible values: SPECIFIC and OTHER.

1. SPECIFIC: the relation has to be between two specific arguments. For an argument to be specific, it must pass either one of the following tests:
  - Unique entity test: the argument must refer to a particular and unique entity.
  - "the" test: the argument is already modified by the determiner "the" or can be paraphrased using the determiner "the" (or the equivalent in a non-English language) without changing its meaning (denotation).

Note it is irrelevant whether the entity's identity is mentioned, or whether it exists in the real world.

2. OTHER: this indicates at least one argument of the relation is not specific.

Examples:

- "John's [lawyer]/E1 [said]/E2 ...": E1 agentOf E2 is SPECIFIC (even though her/his name is not mentioned; "lawyer" passed the unique entity test).
- "Any [ship]/E1 passing through this [region]/E2 is in danger": E1 locatedAt E2 is OTHER.
- "A [ship]/E1 outside of [Somalia coast]/E2 sent out a distress signal": E1 locatedAt E2 is SPECIFIC ("ship" passed the unique entity test).
- "No sane [person]/E1 would [say]/E2 such a thing": E1 agentOf E2 is OTHER.
- "If [he]/E1 wrote the [book]/E2, ...": E1 authourOf E2 is SPECIFIC ("book" passed both tests).
- "[Governments]/E1 around the [world]/E2": E1 locatedAt E2 is OTHER.
- "the [Asian]/E1 [governments]/E2", "the 3 [Asian]/E1 [governments]/E2": E1 locatedAt E2 is SPECIFIC ("governments" passed the "the" test).
- "[Asian]/E1 [governments]/E2": E1 locatedAt E2 is OTHER.
- "for those of [you]/E1 who live in [cities]/E2" (in a speech): E1 residesIn E2 is OTHER.

Comparing tense and modality: These are two different concepts, do not confuse one with the other. For example, if modality of a relation is OTHER (i.e., it is not asserted), it does not imply its tense is UNSPECIFIED ("[He]/E1 was a suspect of the [murder]/E2: E1 agentOf E2 - modality is OTHER but tense is PAST).

### 5.1.7 Anchor words of Relations

For each relation annotated we may want to mark its anchor words - these are the words that trigger the creation of the relation:

1. Anchor words can be multiple words; e.g., phrasal verbs such as "called for".
2. Anchor words are optional, e.g., when anchor words are one of the relation arguments, do not mark it.
3. Anchor words have to strongly suggest the existence of the relation even in a different context; e.g., the word "is" in "he is 27" is not a good anchor word as it is too generic.

## 5.2 Definition of Relation Types

### 5.2.1 affectedBy (tense=UNSPECIFIED)

This relation is between an entity, and an event which has clear directionality, in which the entity is affected by the event.

Examples:

- "[He]/PERSON has been [indicted]/EVENT\_LEGAL.": "he" affectedBy "indicted".
- "Around 300 [people]/PEOPLE were [injured]/EVENT\_VIOLENCE.": "people" affectedBy "injured".

### Argument Type Restrictions

- arg1 = FACILITY | GEOLOGICALOBJ | GPE | LOCATION | ORGANIZATION | PEOPLE | PERSON | PERSONPEOPLE | WEB, arg2 = EVENT | EVENT\_AWARD | EVENT\_BUSINESS | EVENT\_COMMUNICATION | EVENT\_CRIME | EVENT\_CUSTODY | EVENT\_DEMONSTRATION | EVENT\_DISASTER | EVENT\_EDUCATION | EVENT\_ELECTION | EVENT\_LEGAL | EVENT\_LEGISLATION | EVENT\_MEETING | EVENT\_PERFORMANCE | EVENT\_PERSONNEL | EVENT\_SPORTS | EVENT\_VIOLENCE
- arg1 = VEHICLE, arg2 = EVENT | EVENT\_CRIME | EVENT\_DISASTER | EVENT\_VIOLENCE
- arg1 = ANIMAL, arg2 = EVENT | EVENT\_CRIME | EVENT\_CUSTODY | EVENT\_DISASTER | EVENT\_VIOLENCE

### 5.2.2 affiliatedWith (symmetric)

This relation indicates an entity has affiliation with or is connected to another entity:

1. The affiliation can be ethnic, political, religious, or ideological.
2. If there is a more specific relation between the two entities such as affectedBy, citizenOf, clientOf, colleague, educatedAt, employedBy, founderOf, managerOf, memberOf, ownerOf, parentOf, partner, partOf, relative, shareholdersOf, siblingOf, spokespersonFor, spouseOf and subsidiaryOf, use that instead of this relation.

Examples:

- "[he]/PERSON is an [African]/GPE/PEOPLE [American]/GPE/PERSON": "he" coref'ed "American", "American" affiliatedWith "African".
- "[African]/GPE/PEOPLE [Americans]/GPE/PEOPLE today accounted for x% of population...": "Americans" affiliatedWith "African"
- "Abu Sayyaf [Muslim]/PEOPLE [extremist group]/ORGANIZATION": "extremist group" affiliatedWith "Muslim".

## Argument Type Restrictions

- arg1 = GPE | ORGANIZATION | PEOPLE | PERSON | PERSONPEOPLE, arg2 = GPE | ORGANIZATION | PEOPLE | PERSON | PERSONPEOPLE

### 5.2.3 agentOf (tense=UNSPECIFIED)

This relation is between an entity and an event, in which the agent entity plays the most active role in the event according to the text (no background knowledge should be required). If more than one candidate is available for the agent argument, pick the one that has the most direct involvement of the event. E.g., in "Israel sent [ground forces]/PEOPLE to [knock down]/EVENT\_VIOLENCE key elements in Gaza..." we mark "ground forces" agentOf "knock down".

Examples:

- "Two [people]/PEOPLE [shot]/EVENT\_VIOLENCE a leading Turkish human-rights activist.": "people" agentOf "shot"
- "[Hsu Lu]/PERSON [said]/EVENT\_COMMUNICATION that her news staff was reporting that visibility was bad.": "Hsu Lu" agentOf "said".

## Argument Type Restrictions

- arg1 = GPE | ORGANIZATION | PEOPLE | PERSON | PERSONPEOPLE | WEB, arg2 = EVENT | EVENT\_AWARD | EVENT\_BUSINESS | EVENT\_COMMUNICATION | EVENT\_CRIME | EVENT\_CUSTODY | EVENT\_DEMONSTRATION | EVENT\_DISASTER | EVENT\_EDUCATION | EVENT\_ELECTION | EVENT\_LEGAL | EVENT\_LEGISLATION | EVENT\_MEETING | EVENT\_PERFORMANCE | EVENT\_PERSONNEL | EVENT\_SPORTS | EVENT\_VIOLENCE
- arg1 = ANIMAL, arg2 = EVENT | EVENT\_CRIME | EVENT\_VIOLENCE

### 5.2.4 authorOf (tense=UNSPECIFIED)

This relation is between a person and the TITLEWORK s/he created. This relation holds between writer/composer etc and her/his work.

Example:

- "[James Herriot]/PERSON 's [All Creatures Great and Small]/TITLEWORK" is a wonderful book.": "James Herriot" authorOf "All Creatures Great and Small"

## Argument Type Restrictions

- arg1 = ORGANIZATION | PEOPLE | PERSON | PERSONPEOPLE, arg2 = TITLEWORK

### 5.2.5 awardedBy

This relation is between an award or degree that is/was awarded by a person or organization. Note if an actual event is mentioned, it should also be marked by EVENT\_AWARD.

Examples:

- "He was awarded the [National Medal of Technology]/AWARD from President [Ronald Reagan]/PERSON in 1985.": "National Medal of Technology" awardedBy "Ronald Reagan".
- "After graduating with a [Juris Doctor]/DEGREE magna cum laude from [Harvard]/ORGANIZATION in 1991, he returned to Chicago.": "Juris Doctor" awardedBy "Harvard".

## Argument Type Restrictions

- arg1 = AWARD | DEGREE, arg2 = GPE | ORGANIZATION | PEOPLE | PERSON | PERSONPEOPLE



### 5.2.6 awardedTo

This relation is between an award or degree that is/was awarded to a person or organization. Note if an actual event is mentioned, it should also be marked by `EVENT_AWARD`.

Examples:

- "[He]/PERSON was awarded the [National Medal of Technology]/AWARD from President Ronald Reagan in 1985.": "National Medal of Technology" awardedTo "He".
- "After graduating with a [Juris Doctor]/DEGREE magna cum laude from Harvard in 1991, [he]/PERSON returned to Chicago.": "Juris Doctor" awardedTo "he".
- "the [Best Picture]/AWARD was given to the [Godfather]/TITLEWORK...": "Best Picture" awardedTo "Godfather".

#### Argument Type Restrictions

- arg1 = DEGREE, arg2 = GPE | ORGANIZATION | PEOPLE | PERSON | PERSONPEOPLE
- arg1 = AWARD, arg2 = GPE | ORGANIZATION | PEOPLE | PERSON | PERSONPEOPLE | TITLEWORK

### 5.2.7 basedIn

This relation is between an organization and a location in which it is mainly, only, or intrinsically located. E.g., headquarter location of a company.

Examples:

- "[Kent University]/ORGANIZATION here in [Britain]/GPE/LOCATION": "Kent University" basedIn "Britain"
- "Young and Rubican, a [New York]/GPE/LOCATION [PR firm]/ORGANIZATION": "PR firm" basedIn "New York"

#### Argument Type Restrictions

- arg1 = ORGANIZATION, arg2 = FACILITY | GEOLOGICALOBJ | GPE | LOCATION

### 5.2.8 before (transitive, tense=UNSPECIFIED)

This relation signifies the temporal relation before between two times/events. Mark only when the text clearly specifies the relation, e.g., the use of words such as "before", "after", "next", "then", "afterwards" etc.

Examples:

- "the [explosion]/EVENT\_VIOLENCE came after the plane [crashed]/EVENT\_DISASTER into the ground.": "crashed" before "explosion"
- "the explosion happened [two days after]/DATE the first [explosion]/EVENT\_VIOLENCE": "explosion" before "two days after" "two days after" timeOf "explosion"

#### Argument Type Restrictions

- arg1 = DATE | DURATION | EVENT | EVENT\_AWARD | EVENT\_BUSINESS | EVENT\_COMMUNICATION | EVENT\_CRIME | EVENT\_CUSTODY | EVENT\_DEMONSTRATION | EVENT\_DISASTER | EVENT\_EDUCATION | EVENT\_ELECTION | EVENT\_LEGAL | EVENT\_LEGISLATION | EVENT\_MEETING | EVENT\_PERFORMANCE | EVENT\_PERSONNEL | EVENT\_SPORTS | EVENT\_VIOLENCE | TIME, arg2 = DATE | DURATION | EVENT | EVENT\_AWARD | EVENT\_BUSINESS | EVENT\_COMMUNICATION | EVENT\_CRIME | EVENT\_CUSTODY | EVENT\_DEMONSTRATION | EVENT\_DISASTER | EVENT\_EDUCATION | EVENT\_ELECTION | EVENT\_LEGAL | EVENT\_LEGISLATION | EVENT\_MEETING | EVENT\_PERFORMANCE | EVENT\_PERSONNEL | EVENT\_SPORTS | EVENT\_VIOLENCE | TIME

### 5.2.9 bornAt (tense=UNSPECIFIED)

This relation is between a person/animal and a place where s/he/it was born at (birth place).

Example:

- "[Albert Gore, Jr.]/PERSON was born in [Washington, D.C]/GPE/LOCATION": "Albert Gore, Jr." bornAt "Washington, D.C."

#### Argument Type Restrictions

- arg1 = ANIMAL | PEOPLE | PERSON | PERSONPEOPLE, arg2 = FACILITY | GEOLOGICALOBJ | GPE | LOCATION

### 5.2.10 bornOn (tense=UNSPECIFIED)

This relation is between a person/animal and her/his/its date of birth.

Example:

- "My [mom]/PERSON was born in [1919]/DATE.": "mom" bornOn "1919"

#### Argument Type Restrictions

- arg1 = ANIMAL | PEOPLE | PERSON | PERSONPEOPLE, arg2 = DATE | DURATION | TIME

### 5.2.11 capitalOf

This relation is between the capital and its country/state/province. Mark only when the text says so explicitly - do not use world knowledge.

Example:

- "[Washington DC]/GPE/LOCATION is the capital of the [US]/GPE/LOCATION.": "Washington DC" capitalOf "US".

#### Argument Type Restrictions

- arg1 = GPE | LOCATION, arg2 = GPE

### 5.2.12 citizenOf

This relation is between a person and his/her country. Note that is stronger than residence, because it has legal implications.

Example:

- "[Maria Zuchowicz]/PERSON of [Poland]/GPE/LOCATION": "Maria Zuchowicz" citizenOf "Poland".

#### Argument Type Restrictions

- arg1 = PEOPLE | PERSON | PERSONPEOPLE, arg2 = GPE

### 5.2.13 clientOf

This relation is between two entities where one is a direct business client of the other.

Examples:

- "President [Clinton]/PERSON 's personal [lawyer]/PERSON, [David Kendall]/PERSON": "Clinton" clientOf "lawyer" (with the last two PERSONs coref'ed).
- "[Arnold]/PERSON handles advertising for such major [corporations]/ORGANIZATION as ...": "Arnold" clientOf "corporations".
- Negative example: "The [company]/ORGANIZATION indirectly supply many [restaurants]/FACILITY with the peanut-related products": indirect clients should not be marked.

### Argument Type Restrictions

- arg1 = ANIMAL | GPE | PEOPLE | PERSON | PERSONPEOPLE, arg2 = FACILITY | GPE | ORGANIZATION | PEOPLE | PERSON | PERSONPEOPLE
- arg1 = ORGANIZATION, arg2 = GPE | ORGANIZATION | PEOPLE | PERSON | PERSONPEOPLE

#### 5.2.14 colleague (symmetric)

This relation is between two people that are part of the same organization.

Example:

- "[Bakuyev]/PERSON worked closely with [Shamil Bassayev]/PERSON": "Bakuyev" colleague "Shamil Bassayev".

### Argument Type Restrictions

- arg1 = PEOPLE | PERSON | PERSONPEOPLE, arg2 = PEOPLE | PERSON | PERSONPEOPLE

#### 5.2.15 competitor (symmetric)

This relation is between two organizations that are engaged in economic competition.

Example:

- "The rapid expansion of [Goldman]/ORGANIZATION 's chief [competitors]/ORGANIZATION": "Goldman" competitor "competitors".

### Argument Type Restrictions

- arg1 = GPE | ORGANIZATION, arg2 = GPE | ORGANIZATION

#### 5.2.16 diedAt (tense=UNSPECIFIED)

This relation is between a person/animal and a place where s/he/it died (death place).

Example:

- "[Edward Baker Lincoln]/PERSON died on February 1, 1850 in [Springfield]/GPE/LOCATION.": "Edward Baker Lincoln" diedAt "Springfield".

### Argument Type Restrictions

- arg1 = ANIMAL | PEOPLE | PERSON | PERSONPEOPLE, arg2 = FACILITY | GEOLOGICALOBJ | GPE | LOCATION

#### 5.2.17 diedOf (tense=UNSPECIFIED)

This relation is between a person/animal and his/her/its cause of death. This includes diseases, substances (drugs etc), disasters, violent events, and the other events that can be said to be the cause of the death.

Examples:

- "[He]/PERSON died of [cancer]/DISEASE.": "he" diedOf "cancer".
- "[He]/PERSON was killed in [WWII]/EVENT\_VIOLENCE.": "he" diedOf "WWII".

### Argument Type Restrictions

- arg1 = ANIMAL | PEOPLE | PERSON | PERSONPEOPLE, arg2 = DISEASE | EVENT | EVENT\_CRIME | EVENT\_DISASTER | EVENT\_LEGAL | EVENT\_VIOLENCE | SUBSTANCE

### 5.2.18 diedOn (tense=UNSPECIFIED)

This relation is between a person/animal and her/his/its date of death.

Example:

- "[He]/PERSON was murdered [last year]/DATE.": "he" diedOn "last year".

#### Argument Type Restrictions

- arg1 = ANIMAL | PEOPLE | PERSON | PERSONPEOPLE, arg2 = DATE | DURATION | TIME

### 5.2.19 dissolvedOn (tense=UNSPECIFIED)

This relation is between an organization and the date when it was/will be dissolved.

Examples:

- "[Compaq]/ORGANIZATION existed as an independent corporation until [2002]/DATE, when it merged with Hewlett-Packard.": "Compaq" dissolvedOn "2002".
- "[League of Nations]/ORGANIZATION was officially disbanded in [1946]/DATE when the WWII ended." "League of Nations" dissolvedOn "1946".

#### Argument Type Restrictions

- arg1 = GPE | ORGANIZATION | WEB, arg2 = DATE | DURATION | TIME

### 5.2.20 educatedAt (tense=UNSPECIFIED)

This relation is between a person and the organization he receives/received education at (alma mater).

Example:

- "Gore [E1] enrolled in Harvard University [E2] in 1965."

#### Argument Type Restrictions

- arg1 = PEOPLE | PERSON | PERSONPEOPLE, arg2 = ORGANIZATION

### 5.2.21 employedBy

This relation is between two entities where one pays the other for certain work/service. For this relation to hold, the "employer" argument must be represented by a single person/organization, and there must exist monetary reward for the service.

The payment of money does not need to be made directly from the employer entity to the employee entity, as long as there is a parent entity that pays for the service; e.g., in "Supreme Court Justice Ginsburg" we mark Ginsburg employedBy Supreme Court although technically she is paid by the US government for her service.

A corollary of the above is that whenever choices are available between marking employedBy with an organization and its parent organization, always choose the one the employee has the closest affiliation with (see examples below).

Examples:

- "[Jobs]/PERSON was hired by [HP]/ORGANIZATION as a summer employee": "Jobs" employedBy "HP"
- "Clinton was [New York]/GPE/LOCATION [senator]/PERSON": "senator" employedBy "New York"; although technically Clinton represents NY, we decided to mark this representation relationship with employedBy.

- "[Ben]/PERSON works for [IBM Research]/ORGANIZATION": "Ben" employedBy "IBM Research"; although it's IBM (the company) that pays for Ben's work, we mark employedBy relation between Ben and his closest affiliation (and we can always relate IBM Research to IBM via partOf if it's available from the text).
- "former [Army]/ORGANIZATION Captain [James Pietrangelo II]/PERSON": "James Pietrangelo II" employedBy "Army"; same as the reasoning above - technically Pietrangelo was paid by the US government, but Army was his closest affiliation; if the text is "US Army captain..." we still mark Pietrangelo employedBy Army, and have Army partOf US.
- Negative example: "[Hollywood]/ORGANIZATION [actors]/PEOPLE" - the relation does not hold because Hollywood is not a single organization; same goes with "Wall Street broker".

### Argument Type Restrictions

- arg1 = GPE | PEOPLE | PERSON | PERSONPEOPLE, arg2 = GPE | ORGANIZATION | PEOPLE | PERSON | PERSONPEOPLE | WEB

#### 5.2.22 foundedOn (tense=UNSPECIFIED)

This relation is between an organization and the time it was founded at.

Example:

- "[IBM]/ORGANIZATION was founded in [1896]/DATE": "IBM" foundedOn "1896".

### Argument Type Restrictions

- arg1 = GPE | ORGANIZATION | WEB, arg2 = DATE | DURATION | TIME

#### 5.2.23 founderOf (tense=UNSPECIFIED)

This relation is between a person and the organization he started.

Example:

- "the [founder]/PERSON of the [First Commercial Bank]/ORGANIZATION in Lublin": "founder" founderOf "First Commercial Bank".

### Argument Type Restrictions

- arg1 = PEOPLE | PERSON | PERSONPEOPLE, arg2 = FACILITY | GPE | ORGANIZATION | WEB

#### 5.2.24 hasAttribute

This relation is between an entity and an attribute it has, such as age, disease, height/weight, email address, phone/fax numbers, websites, stock ticker symbols, and money (see below).

If the second argument is MONEY, it must fall under one of these categories: net worth, revenue, budget, income, profit, earnings, loss.

Examples:

- "[She]/PERSON was [5 years old]/AGE": "she" hasAttribute "5 years old".
- "You can reach [him]/PERSON at [john@foo.com]/EMAIL": "him" hasAttribute "john@foo.com".
- "Microsoft Chairman [Bill Gates]/PERSON has had his personal net worth soar [over 100 billion dollars]/MONEY": "Bill Gates" hasAttribute "over 100 billion dollars".
- "[IBM]/ORGANIZATION reported second-quarter revenue of [\$23.8bn]/MONEY": "IBM" hasAttribute "\$23.8bn".
- "[Microsoft Corp.]/ORGANIZATION ([MSFT]/TICKER) gave up 64 cents, or 3.6%, to \$17.48": "Microsoft" hasAttribute "MSFT".
- Negative example: "[He]/PERSON earned an [MBA]/DEGREE from Stanford" - use awardedTo instead.

### Argument Type Restrictions

- arg1 = GPE | ORGANIZATION, arg2 = AGE | EMAIL | MONEY | PHONE | TICKER | WEB
- arg1 = PEOPLE | PERSON | PERSONPEOPLE, arg2 = AGE | DISEASE | EMAIL | MEASURE | MONEY | PHONE | WEB
- arg1 = FACILITY | GEOLOGICALOBJ | LOCATION | PLANT | PRODUCT | SUBSTANCE | VEHICLE | WEAPON, arg2 = AGE
- arg1 = ANIMAL, arg2 = AGE | DISEASE

#### 5.2.25 instrumentOf (tense=UNSPECIFIED)

This relation is between an entity and an event that the entity is used to bring about.

Examples:

- "I [killed]/EVENT\_VIOLENCE him with my [gun]/WEAPON": "gun" instrumentOf "killed".
- "The 9/11 [attacks]/EVENT\_VIOLENCE were carried out using four [jumbo jets]/VEHICLE": "jumbo jets" instrumentOf "attacks"
- "the [swine flue]/DISEASE [outbreak]/EVENT\_DISASTER": "swine flue" instrumentOf "outbreak".

### Argument Type Restrictions

- arg1 = SUBSTANCE | VEHICLE, arg2 = EVENT\_CRIME | EVENT\_VIOLENCE
- arg1 = DISEASE, arg2 = EVENT\_CRIME | EVENT\_DISASTER
- arg1 = WEAPON, arg2 = EVENT | EVENT\_CRIME | EVENT\_VIOLENCE

#### 5.2.26 locatedAt (transitive)

This relation is between an entity and its location.

1. If both arguments of the relation are locations, pick the one that has a smaller geographical footprint as the first argument.
2. For the argument 2 (the location), if more than one candidate are available, choose the one that has a smaller geographical footprint. E.g., in "[Haniya] [preached] in a [Gaza] [mosque]" we mark "preached" locatedAt "mosque" and "mosque" locatedAt "Gaza" (together with "Haniya" agentOf "preached").
3. In contrast to basedIn, residesIn, and citizenOf, locatedAt does not imply any information other than the entity's physical whereabouts.

Examples:

- "[Arafat]/PERSON is coming to [Washington]/GPE/LOCATION": "Arafat" locatedAt "Washington".
- "An Israeli [military court]/ORGANIZATION in [Lod]/GPE/LOCATION": "military court" locatedAt "Lod".

### Argument Type Restrictions

- arg1 = ANIMAL | PEOPLE | PERSON | PERSONPEOPLE | WEAPON, arg2 = FACILITY | GEOLOGICALOBJ | GPE | LOCATION | VEHICLE
- arg1 = SUBSTANCE, arg2 = FACILITY | GEOLOGICALOBJ | GPE | LOCATION | VEHICLE | WEAPON
- arg1 = EVENT | EVENT\_AWARD | EVENT\_BUSINESS | EVENT\_COMMUNICATION | EVENT\_CRIME | EVENT\_CUSTODY | EVENT\_DEMONSTRATION | EVENT\_DISASTER | EVENT\_EDUCATION | EVENT\_ELECTION | EVENT\_LEGAL | EVENT\_LEGISLATION | EVENT\_MEETING | EVENT\_PERFORMANCE | EVENT\_PERSONNEL | EVENT\_SPORTS | EVENT\_VIOLENCE | FACILITY | GEOLOGICALOBJ | GPE | LOCATION | ORGANIZATION | PLANT | VEHICLE, arg2 = FACILITY | GEOLOGICALOBJ | GPE | LOCATION

#### 5.2.27 managerOf

This relation is between a person and another entity that s/he manages as her/his job.

Examples:

- "[Iraq]/GPE 's [president]/PERSON [Saddam Hussein]/PERSON": two PERSONs are coref'ed and "president" managerOf "Iraq".
- "a former [mayor]/PERSON of [Shanghai]/GPE/LOCATION": "mayor" managerOf "Shanghai".

### Argument Type Restrictions

- arg1 = PEOPLE | PERSON | PERSONPEOPLE, arg2 = FACILITY | GPE | LOCATION | ORGANIZATION | PEOPLE | PERSON | PERSONPEOPLE | WEB

#### 5.2.28 memberOf

This relation is between an entity and another entity to which the former entity belongs.

Use the relation managerOf if the argument mentions apply. E.g., in "[Steve Jobs]/PERSON, the [CEO]/PERSON of [Apple]/ORGANIZATION, ...", mark "CEO" managerOf "Apple" instead of "CEO" memberOf "Apple" and "CEO" employedBy "Apple".

Examples:

- "The [Lakers]/ORGANIZATION ' [Nick Van Exel]/PERSON": "Nick Van Exel" memberOf "Lakers".
- "[West Germany]/GPE became a member of [NATO]/ORGANIZATION on 9 May 1955": "West Germany" memberOf "NATO".
- Negative example: "[Frank]/PERSON is [Irish]/GPE/PEOPLE": "Frank" partOfMany "Irish".

### Argument Type Restrictions

- arg1 = PEOPLE | PERSON | PERSONPEOPLE, arg2 = FACILITY | GPE | ORGANIZATION | WEB
- arg1 = GPE | ORGANIZATION, arg2 = GPE | ORGANIZATION

### Comparisons between memberOf, partOf and partOfMany:

#### 1. memberOf vs partOf

(a) memberOf is not transitive but partOf is; examples:

- John memberOf SONY and SONY memberOf RIAA does not imply John memberOf RIAA

- but Salim's group partOf IBM Research and IBM Research partOf IBM does imply Salim's group partOf IBM.
- (b) memberOf implies autonomy, while partOf does not; examples:
- US memberOf NATO (not US partOf NATO) because US is an independent country
  - France memberOf EU (because the EU is a loose federation)
  - Canada memberOf Commonwealth (despite having the Queen of England as its head of state)
  - But: Appenzell Innerrhoden partOf Switzerland (because the Swiss federation is a tight organization);

## 2. memberOf vs partOfMany

- (a) memberOf is not transitive but partOfMany is.
- (b) For two arguments related by partOfMany, the two must be of the same or closely related type - this is not necessarily true for memberOf relation (unless both arguments are organizations/company roles).
- (c) Also note partOfMany indicates a quantitative relation between the two arguments (the first argument is a strict subset of the second argument), and it does not indicate a qualitative difference between the two (e.g., Frank partOfMany Irish indicates no difference in terms of qualities of Frank and all the Irish people - they're all Irish); but memberOf does imply a qualitative difference between the two arguments involved.

### 5.2.29 near (symmetric)

This relation is between any pair of entities that are physically located near to each other.

Example:

- "a [mansion]/FACILITY overlooking the [Pacific Ocean]/GEOLOGICALOBJ": "mansion" near "Pacific Ocean".

### Argument Type Restrictions

- arg1 = ANIMAL | PEOPLE | PERSON | PERSONPEOPLE | WEAPON, arg2 = FACILITY | GEOLOGICALOBJ | GPE | LOCATION | VEHICLE
- arg1 = SUBSTANCE, arg2 = FACILITY | GEOLOGICALOBJ | GPE | LOCATION | VEHICLE | WEAPON
- arg1 = EVENT | EVENT\_AWARD | EVENT\_BUSINESS | EVENT\_COMMUNICATION | EVENT\_CRIME | EVENT\_CUSTODY | EVENT\_DEMONSTRATION | EVENT\_DISASTER | EVENT\_EDUCATION | EVENT\_ELECTION | EVENT\_LEGAL | EVENT\_LEGISLATION | EVENT\_MEETING | EVENT\_PERFORMANCE | EVENT\_PERSONNEL | EVENT\_SPORTS | EVENT\_VIOLENCE | FACILITY | GEOLOGICALOBJ | GPE | LOCATION | ORGANIZATION | PLANT | VEHICLE, arg2 = FACILITY | GEOLOGICALOBJ | GPE | LOCATION

### 5.2.30 overlaps (symmetric, tense=UNSPECIFIED)

This relation signifies the temporal relation overlaps between two times/events; i.e., the time of the first argument overlaps with the time of the second argument. Mark only when the text clearly specifies the relation, e.g., the use of words such as "when", "meanwhile", "at the same time" etc.

Example:

- "There was an [explosion]/EVENT\_VIOLENCE as the plane [crashed]/EVENT\_DISASTER into the ground": "explosion" overlaps "crashed"



## Argument Type Restrictions

- arg1 = EVENT | EVENT\_AWARD | EVENT\_BUSINESS | EVENT\_COMMUNICATION | EVENT\_CRIME | EVENT\_CUSTODY | EVENT\_DEMONSTRATION | EVENT\_DISASTER | EVENT\_EDUCATION | EVENT\_ELECTION | EVENT\_LEGAL | EVENT\_LEGISLATION | EVENT\_MEETING | EVENT\_PERFORMANCE | EVENT\_PERSONNEL | EVENT\_SPORTS | EVENT\_VIOLENCE, arg2 = EVENT | EVENT\_AWARD | EVENT\_BUSINESS | EVENT\_COMMUNICATION | EVENT\_CRIME | EVENT\_CUSTODY | EVENT\_DEMONSTRATION | EVENT\_DISASTER | EVENT\_EDUCATION | EVENT\_ELECTION | EVENT\_LEGAL | EVENT\_LEGISLATION | EVENT\_MEETING | EVENT\_PERFORMANCE | EVENT\_PERSONNEL | EVENT\_SPORTS | EVENT\_VIOLENCE

### 5.2.31 ownerOf

This relation is between an owner and an entity it/she/he owns. The possession relationship does not need to be permanent.

Examples:

- "Spotty is [her]/PERSON [cat]/ANIMAL": "her" ownerOf "cat".
- "It was identified as the [suspect]/PERSON 's [car]/VEHICLE": "suspect" ownerOf "car" (although the car might be stolen or rented).
- "[His]/PERSON [apartment]/FACILITY": "his" ownerOf "apartment" (although he might be renting).

## Argument Type Restrictions

- arg1 = GPE | ORGANIZATION | PEOPLE | PERSON | PERSONPEOPLE, arg2 = ANIMAL | FACILITY | LOCATION | ORGANIZATION | VEHICLE | WEAPON | WEB

### 5.2.32 parentOf (tense=UNSPECIFIED)

This relation is between a parent and her/his child. Include step-parents/children, but do not include in-laws.

Example:

- "his father passed away when he was 17": "father" parentOf "his".

## Argument Type Restrictions

- arg1 = ANIMAL, arg2 = ANIMAL
- arg1 = PEOPLE | PERSON | PERSONPEOPLE, arg2 = PEOPLE | PERSON | PERSONPEOPLE

### 5.2.33 participantIn (tense=UNSPECIFIED)

This relation is between a participant and an event she/he/it participated/participates. If the participant is the agent of the event or can be interpreted as being affected by the event, use agentOf and affectedBy instead, respectively.

Example:

- "The [President]/PERSON and the visiting [leaders]/PEOPLE [discussed]/EVENT\_MEETING Islamic activities": "president" participantIn "discussed" and "leaders" participantIn "discussed".

## Argument Type Restrictions

- arg1 = GPE | ORGANIZATION | PEOPLE | PERSON | PERSONPEOPLE, arg2 = EVENT | EVENT\_AWARD | EVENT\_BUSINESS | EVENT\_COMMUNICATION | EVENT\_CRIME | EVENT\_CUSTODY | EVENT\_DEMONSTRATION | EVENT\_DISASTER | EVENT\_EDUCATION | EVENT\_ELECTION | EVENT\_LEGAL | EVENT\_LEGISLATION | EVENT\_MEETING | EVENT\_PERFORMANCE | EVENT\_PERSONNEL | EVENT\_SPORTS | EVENT\_VIOLENCE
- arg1 = ANIMAL, arg2 = EVENT | EVENT\_CRIME | EVENT\_PERFORMANCE | EVENT\_VIOLENCE

### 5.2.34 partner (symmetric)

This relation is between two organizations that are engaged in economic cooperation.

Example:

- "...two seed companies that had been [its]/ORGANIZATION [partners]/ORGANIZATION in creating genetically modified crops": "its" partner "partners".

## Argument Type Restrictions

- arg1 = GPE | ORGANIZATION, arg2 = GPE | ORGANIZATION

### 5.2.35 partOf (transitive)

This relation is between smaller and larger entities of the same or related type, where the second entity subsumes the first.

If the two arguments are events, the smaller event must take place within the time span of the larger event for these two events to be related by this relation.

This relation is tenseless (tense=UNSPECIFIED) when both of the arguments are events.

Examples:

- "Arafat's [Fatah faction]/ORGANIZATION of the [PLO]/ORGANIZATION": "Fatah faction" partOf "PLO".
- "The [attacks]/EVENT\_VIOLENCE are part of an escalating arson and bombing [spree]/EVENT\_VIOLENCE in recent months": "attacks" partOf "spree".

## Argument Type Restrictions

- arg1 = EVENT | EVENT\_AWARD | EVENT\_BUSINESS | EVENT\_COMMUNICATION | EVENT\_CRIME | EVENT\_CUSTODY | EVENT\_DEMONSTRATION | EVENT\_DISASTER | EVENT\_EDUCATION | EVENT\_ELECTION | EVENT\_LEGAL | EVENT\_LEGISLATION | EVENT\_MEETING | EVENT\_PERFORMANCE | EVENT\_PERSONNEL | EVENT\_SPORTS | EVENT\_VIOLENCE, arg2 = EVENT | EVENT\_AWARD | EVENT\_BUSINESS | EVENT\_COMMUNICATION | EVENT\_CRIME | EVENT\_CUSTODY | EVENT\_DEMONSTRATION | EVENT\_DISASTER | EVENT\_EDUCATION | EVENT\_ELECTION | EVENT\_LEGAL | EVENT\_LEGISLATION | EVENT\_MEETING | EVENT\_PERFORMANCE | EVENT\_PERSONNEL | EVENT\_SPORTS | EVENT\_VIOLENCE
- arg1 = GPE | ORGANIZATION, arg2 = GPE | ORGANIZATION

## Comparisons between partOf, memberOf and partOfMany:

1. partOf vs memberOf: See Section 5.2.28 on memberOf vs partOf.
2. partOf vs partOfMany: for partOfMany the elements of both of the arguments have no qualitative difference, but the same cannot be said for partOf. E.g., in "Frank is Irish" (Frank partOfMany Irish), Frank is an Irish, but in IBM Research partOf IBM, IBM Research is not IBM. partOfMany

### 5.2.36 (transitive, tense=UNSPECIFIED)

This relation is between two entities of the same or related type: if E1 partOfMany E2, then E2 is plural and E1 is one or more elements in E2.

Examples:

- "The 12,000 [people]/PEOPLE attending the conference included [scientists]/PEOPLE and [doctors]/PEOPLE": "scientists" partOfMany "people" and "doctors" partOfMany "people".
- "[Three]/CARDINAL/PEOPLE of the five [students]/PEOPLE went home": "three" partOfMany "students".

#### Argument Type Restrictions

- arg1 = PEOPLE | PERSON | PERSONPEOPLE, arg2 = PEOPLE | PERSON | PERSONPEOPLE
- arg1 = DISEASE, arg2 = DISEASE
- arg1 = VEHICLE, arg2 = VEHICLE
- arg1 = EVENT\_EDUCATION, arg2 = EVENT | EVENT\_EDUCATION
- arg1 = EVENT, arg2 = EVENT
- arg1 = GEOLOGICALOBJ, arg2 = GEOLOGICALOBJ
- arg1 = EVENT\_PERSONNEL, arg2 = EVENT | EVENT\_PERSONNEL
- arg1 = GPE | ORGANIZATION, arg2 = GPE | ORGANIZATION
- arg1 = LOCATION, arg2 = LOCATION
- arg1 = TITLEWORK, arg2 = TITLEWORK
- arg1 = EVENT\_LEGISLATION, arg2 = EVENT | EVENT\_LEGISLATION
- arg1 = EVENT\_COMMUNICATION, arg2 = EVENT | EVENT\_COMMUNICATION
- arg1 = EVENT\_ELECTION, arg2 = EVENT | EVENT\_ELECTION
- arg1 = WEAPON, arg2 = WEAPON arg1 = EVENT\_DISASTER, arg2 = EVENT | EVENT\_DISASTER
- arg1 = EVENT\_PERFORMANCE, arg2 = EVENT | EVENT\_PERFORMANCE
- arg1 = EVENT\_VIOLENCE, arg2 = EVENT | EVENT\_VIOLENCE
- arg1 = EVENT\_AWARD, arg2 = EVENT | EVENT\_AWARD
- arg1 = LAW, arg2 = LAW arg1 = EVENT\_CRIME, arg2 = EVENT | EVENT\_CRIME
- arg1 = FACILITY, arg2 = FACILITY
- arg1 = EVENT\_DEMONSTRATION, arg2 = EVENT | EVENT\_DEMONSTRATION
- arg1 = EVENT\_CUSTODY, arg2 = EVENT | EVENT\_CUSTODY
- arg1 = EVENT\_BUSINESS, arg2 = EVENT | EVENT\_BUSINESS
- arg1 = EVENT\_MEETING, arg2 = EVENT | EVENT\_MEETING
- arg1 = EVENT\_LEGAL, arg2 = EVENT | EVENT\_LEGAL
- arg1 = ANIMAL, arg2 = ANIMAL
- arg1 = EVENT\_SPORTS, arg2 = EVENT | EVENT\_SPORTS

## Comparisons between partOfMany, memberOf and partOf:

1. partOfMany vs memberOf: See Section 5.2.28 on memberOf vs partOfMany.
2. partOfMany vs partOf: See Section 5.2.35 on partOf vs partOfMany.

### 5.2.37 playsRoleOf

This relation is between a person and the character she/he plays/played in a drama, ballet, etc.

Do not use this relation between the person and the function of her/his role; e.g., in "she is the main character in the movie", "she" and "main character" should be coref'ed, not related by this relation.

A side note: DO NOT coref a person and the role s/he plays (see Section 4 on coreference).

Examples:

- "[She]/PERSON plays "[Juliet]/PERSON" in this production": "she" playsRoleOf "Juliet".
- "[Mary]/PERSON and [Sue]/PERSON played the [smart women]/PEOPLE": "Mary" playsRoleOf "smart women", "Sue" playsRoleOf "smart women".

### Argument Type Restrictions

- arg1 = PEOPLE | PERSON | PERSONPEOPLE, arg2 = PEOPLE | PERSON | PERSONPEOPLE

### 5.2.38 populationOf

This relation is between a place and the number of people located there, or a company/organization and the number of employees/members it has. Note this applies to sub-organizations (such as boards, executive teams etc) as well.

Examples:

- "The population of [New York City]/GPE/LOCATION was [8 million]/CARDINAL as of April 2000.": "8 million" populationOf "New York City".
- "With over [388,000]/CARDINAL employees worldwide, [IBM]/ORGANIZATION is the largest IT company in the world": "388,000" populationOf "IBM".
- "The move was unanimously approved Friday by the society's [25]/CARDINAL -member [executive committee]/ORGANIZATION": "25" populationOf "executive committee".

### Argument Type Restrictions

- arg1 = CARDINAL, arg2 = FACILITY | GEOLOGICALOBJ | GPE | LOCATION | ORGANIZATION

### 5.2.39 productOf

This relation is between a product and its company.

Examples:

- "[Microsoft]/ORGANIZATION [Windows]/PRODUCT": "Windows" productOf "Microsoft".
- "The 2009 [Honda]/ORGANIZATION [CR-V]/VEHICLE/PRODUCT is still America's most-popular compact SUV": "CR-V" productOf "Honda".

### Argument Type Restrictions

- arg1 = PRODUCT, arg2 = ORGANIZATION

#### 5.2.40 relative (symmetric)

This relation is between a person and her/his relatives. Note for specific relatives such as parents, children, siblings, and spouses, please use `parentOf`, `siblingOf` and `spouseOf` instead.

Examples:

- "[my] [mother-in-law]/PERSON": "my" relative "mother-in-law".
- Negative example: "[my]/PERSON [father]/PERSON" - mark it "father" `parentOf` "my" instead.

#### Argument Type Restrictions

- `arg1 = ANIMAL, arg2 = ANIMAL`
- `arg1 = PEOPLE | PERSON | PERSONPEOPLE, arg2 = PEOPLE | PERSON | PERSONPEOPLE`

#### 5.2.41 residesIn

This relation is between a person and the place s/he lived/lives in. It is stronger than `locatedAt`, because it implies permanence.

Examples:

- "[Einstein]/PERSON moved to the [United States]/GPE/LOCATION": "Einstein" `residesIn` "United States".
- "an elderly [Moscow]/GPE/LOCATION [woman]/PERSON": "woman" `residesIn` "Moscow".
- "[artists]/PEOPLE living in [Brooklyn]/GPE/LOCATION": "artists" `residesIn` "Brooklyn".

#### Argument Type Restrictions

- `arg1 = ANIMAL | PEOPLE | PERSON | PERSONPEOPLE, arg2 = FACILITY | GEOLOGICALOBJ | GPE | LOCATION`

#### 5.2.42 shareholdersOf

This relation indicates a person/group of people/organization is/are shareholder(s) of an organization.

Examples:

- "[Standard Life Investments]/ORGANIZATION, a top-ten shareholder, signaled that it would be voting against the discretionary payments at [Shell]/ORGANIZATION 's annual meeting": "Standard Life Investments" `shareholdersOf` "Shell".
- "the company has sent letters to [Hapag-Lloyd]/ORGANIZATION 's large [shareholders]/PEOPLE, including [Deutsche Lufthansa AG]/ORGANIZATION ...": "shareholders" `shareholdersOf` "Hapag-Lloyd", "Deutsche Lufthansa AG" `partOfMany` "shareholders".

#### Argument Type Restrictions

- `arg1 = GPE | ORGANIZATION | PEOPLE | PERSON | PERSONPEOPLE, arg2 = ORGANIZATION`

#### 5.2.43 siblingOf (symmetric, tense=UNSPECIFIED)

This relation is between a person and her/his sibling. Include step-brothers/sisters, but do not include in-laws.

Example:

- "[his]/PERSON [sister]/PERSON didn't come": "his" `siblingOf` "sister".

### Argument Type Restrictions

- arg1 = ANIMAL, arg2 = ANIMAL
- arg1 = PEOPLE | PERSON | PERSONPEOPLE, arg2 = PEOPLE | PERSON | PERSONPEOPLE

#### 5.2.44 spokespersonFor

This relation is between a spokesperson and the entity she/he represents.

1. It should be indicated explicitly from the text that the person in question is a spokesperson.
2. Representatives of organizations are also spokespeople.

Example:

- "a Chinese [foreign ministry]/ORGANIZATION [spokesman]/PERSON [Shen Guofang]/PERSON said": "spokesman" spokespersonFor "foreign ministry" and "spokesman" should be coref'ed with "Shen Guofang".

### Argument Type Restrictions

- arg1 = PEOPLE | PERSON | PERSONPEOPLE, arg2 = FACILITY | GPE | ORGANIZATION | PEOPLE | PERSON | PERSONPEOPLE

#### 5.2.45 spouseOf (symmetric)

This relation is between two spouses (not necessarily heterosexual, as long as they went through an official process).

Example:

- "[his]/PERSON [wife]/PERSON came with him": "his" spouseOf "wife".

### Argument Type Restrictions

- arg1 = ANIMAL, arg2 = ANIMAL
- arg1 = PEOPLE | PERSON | PERSONPEOPLE, arg2 = PEOPLE | PERSON | PERSONPEOPLE

#### 5.2.46 subsidiaryOf

This relation indicates an organization is a subsidiary of another organization. To be a subsidiary, it must have fair amount of autonomy (so it is not just partOf the parent organization), and the parent organization must have control over the subsidiary (so the subsidiary is not just memberOf the parent organization).

Examples:

- "the [Africa Export/Import Bank]/ORGANIZATION, a [subsidiary]/ORGANIZATION of the [African Development Bank]/ORGANIZATION based in Cairo, Egypt": "subsidiary" subsidiaryOf "African Development Bank" and "Africa Export/Import Bank" should be coref'ed with "subsidiary".
- "the parent company of [Philip Morris USA]/ORGANIZATION, [Altria Group, Inc.]/ORGANIZATION, announced that...": "Philip Morris USA" subsidiaryOf "Altria Group, Inc.".

### Argument Type Restrictions

- arg1 = ORGANIZATION, arg2 = ORGANIZATION

### 5.2.47 timeOf (tense=UNSPECIFIED)

This relation relates an event to the time when it took place/started/ended, or a TITLEWORK to the time when it was created/performed/broadcasted, or a LAW to the time when it was drafted/created/passed/abolished.

Example:

- "They [met]/EVENT\_MEETING on [Thursday]/DATE": "Thursday" timeOf "met".

#### Argument Type Restrictions

- arg1 = DATE, arg2 = EVENT | EVENT\_AWARD | EVENT\_BUSINESS | EVENT\_COMMUNICATION | EVENT\_CRIME | EVENT\_CUSTODY | EVENT\_DEMONSTRATION | EVENT\_DISASTER | EVENT\_EDUCATION | EVENT\_ELECTION | EVENT\_LEGAL | EVENT\_LEGISLATION | EVENT\_MEETING | EVENT\_PERFORMANCE | EVENT\_PERSONNEL | EVENT\_SPORTS | EVENT\_VIOLENCE | LAW | TITLEWORK
- arg1 = DURATION | TIME, arg2 = EVENT | EVENT\_AWARD | EVENT\_BUSINESS | EVENT\_COMMUNICATION | EVENT\_CRIME | EVENT\_CUSTODY | EVENT\_DEMONSTRATION | EVENT\_DISASTER | EVENT\_EDUCATION | EVENT\_ELECTION | EVENT\_LEGAL | EVENT\_LEGISLATION | EVENT\_MEETING | EVENT\_PERFORMANCE | EVENT\_PERSONNEL | EVENT\_SPORTS | EVENT\_VIOLENCE | TITLEWORK