

Rich ERE Annotation Guidelines Overview V4.2

**Linguistic Data Consortium
April 15, 2016**

© 2016 Trustees of the University of Pennsylvania

This document is unpublished and intended solely for the use of the individual or entity to whom it was delivered. Redistribution is strictly prohibited without the express authorization of the Linguistic Data Consortium.

1 Overview

This document presents a summary of the Rich ERE Annotation Task. Complete annotation guidelines are distributed with the Rich ERE corpora.

Given the variety of approaches and evaluations within DEFT, we set out to define an annotation task that would be supportive of multiple research directions and evaluations, and that would provide a useful foundation for more specialized annotation tasks like inference and anomaly. The resulting Entities, Relations and Events (ERE) annotation task has evolved over the course of the program, from a fairly lightweight treatment of entities, relations and events in text, to a richer representation of phenomena of interest to the program.

In the first phase of the program, we defined Light ERE as a simplified form of ACE annotation, with the goal of being able to rapidly produce consistently labeled data in multiple languages. In Phase 2, Rich ERE expanded entity, relation and event ontologies and expanded the notion of what is taggable. Rich ERE also introduced the notion of *Event Hopper* to address the pervasive challenge of event co-reference, particularly with respect to event mention and event argument granularity variation within and across documents, thus paving the way for the important goal of creating (hierarchical or nested) cross-document event representations. In 2016, we are adding individuality attributes for specific entities to support downstream annotation such as Entity Linking, in which individual specific entities would be linked to a knowledge base. For 2016's evaluation, the inventory of Event types is reduced to 8 types, 18 subtypes.

The sections that follow summarize the annotation approach for each subtask:

- Section 2: Entities
- Section 3: Argument Fillers
- Section 4: Relations
- Section 5: Events
- Section 6: Data Format

2 Entities

An entity is a unique object or set of objects in the world – for instance, a specific person, place, or organization. A mention is a single occurrence of a name, nominal phrase, or pronominal phrase that refers to, or describes, a single entity. The mention extent is a string of text that we annotate to indicate the occurrence of an entity mention. We link together multiple mentions of the same entity during entity coreference annotation.

Entities may be referenced in text at three different mention levels – a name, a common noun or noun phrase, or a pronoun. This is referred as Mention Level Type. In addition to an entity’s mention level – either NAM, NOM, or PRO – annotators will also decide on an entity’s level of specificity, which is one of the two entity classes: Specific or Non-Specific. Entities that are labeled as Specific are also labelled with individuality features -- Individual, Group, or Unknown.

We label five entity types in Rich ERE. There are no entity subtypes.

- **Person (PER)** - Person entities are limited to humans. A PER entity may be a single person or a group.
- **Organization (ORG)** - Organization entities are corporations, agencies, and other groups of people defined by an established organizational structure. An ORG entity may be a single organization or a group. NOTE: A key feature of an ORG is that it can change members without changing identity.
- **Geopolitical Entity (GPE)** - GPE entities are composite entities, consisting of a physical location, a government, and a population. All three of these elements must be present for an entity to be tagged as a GPE. A GPE entity may be a single geopolitical entity or a group.
- **Location (LOC)** - Location entities are geographical entities such as geographical areas and landmasses, bodies of water, and geological formations. A LOC entity may be a single location or a group.
- **Facility (FAC)** – A facility is a functional, primarily man-made structure. Facilities are artifacts falling under the domains of architecture and civil engineering.

The approach to entities is summarized in Table 1 below.

Entity Types	Entity Mention Levels	Entity Classes	Individuality (Specific entities only)
PER	NAM	specific	individual
ORG	NOM	nonspecific	group
GPE	PRO		unknown
LOC			
FAC			

Table 1. Rich ERE Entity Annotation Categories

3 Argument Fillers

Argument Fillers are entity-like participants in relations and events that are not annotated at the entity level. Argument fillers are annotated only when they fill argument roles in tagged relations or events. Argument fillers include: Title, Age, URL, Sentence, Crime, Money, Vehicle, Weapon, Commodity, and Time types.

Each of these argument fillers corresponds to specific relation or event subtypes, meaning that they will only appear if the corresponding subtype lends itself to such information. For example, a person’s age is only annotated as an argument filler in a generalaffiliation-personage relation, and a weapon is annotated only in a limited number of event subtypes, including Conflict.Attack, Manufacture.Artifact, and Life.Injure.

4 Relations

The goal of the Relation Task is, for a specified set of types and subtypes, to detect and characterize relations between tagged entities and argument fillers that are explicitly mentioned in the document.

Relation types and subtypes labeled in Rich ERE are summarized below. Note that the type and subtype inventory in Rich ERE is significantly expanded as compared with Light ERE, in order to be better-aligned with the TAC KBP Slot Filling evaluation. Table 2 summarizes relation types/subtypes in Light and Rich ERE.

Light		Rich	
Type	Subtype	Subtype	Type
Physical	n/a	orgheadquarter	physical
	Located	locatednear	
		resident	
	origin	orglocationorigin	general affiliation
member-origin-religion-ethnicity (more)			
	n/a	personage	general affiliation
	n/a	orgwebshite	
	n/a	org-political-religious-affiliation (opra)	
part whole	subsidiary	subsidiary	part whole
personal social	membership	membership	personal social
	business	business	
	family	family	
	unspecified	unspecified	
	role	role	
org affiliation	employmentmembership	employmentmembership	org affiliation
	leadership	leadership	

	n/a	investshareholder	
	n/a	studentalum	
	n/a	ownership	
	n/a	founder	

Table 2. ERE Relation Types and Subtypes

Each relation has two arguments; arguments can be populated by the labeled entities or argument fillers described in Sections 2 and 3 above. Each relation has a template which specifies what entities/argument fillers can occupy arg1 and arg2. For instance, Physical.Resident relations arg1 is always populated by the PER resident, and arg2 is always populated by the GPE, LOC or FAC place. Appendix 1 shows allowable arguments for each relation type and subtype.

Relations are also labeled for their realis attribute: ASSERTED or OTHER.

Relation coreference is fully automated. Relation mentions that meet all of the following criteria are automatically co-referred:

- They have the same type and subtype
- They have the same realis attribute
- If relations are asymmetric,
 - o $\text{relation1.arg1} == \text{relation2.arg1}$ and $\text{relation1.arg2} == \text{relation2.arg2}$
- If relations are symmetric
 - o $\text{relation1.arg1} == \text{relation2.arg1}$ and $\text{relation1.arg2} == \text{relation2.arg2}$; OR
 - o $\text{relation1.arg1} == \text{relation2.arg2}$ and $\text{relation1.arg2} == \text{relation2.arg1}$

The following three relation type-subtypes are symmetric:

- personalsocial / business
- personalsocial / family
- personalsocial / unspecified

Relation mentions that have fillers as argument are treated as singletons, because unlike entities, argument fillers are not coreferenced.

5 Events

A taggable event is an explicit occurrence of an event with or without participants. The goal of event annotation is to detect and characterize events that tagged entities and argument fillers participate in, and to perform event co-reference. For each event mention, we label the event type and subtype, its

realis attribute, any of its arguments or participants that are within the event mention scope¹, and a required textual “trigger”.

We tag a limited inventory of event types/subtypes. For 2016’s evaluation, the inventory is further reduced from 9 types, 38 subtypes to 8 types, 18 subtypes, shown in Table 3 below.

type	subtype	2016 eval
business	declareBankruptcy	
business	endOrg	
business	mergeOrg	
business	startOrg	
conflict	attack	yes
conflict	demonstrate	yes
contact	broadcast	yes
contact	contact	yes
contact	correspondence	yes
contact	meet	yes
justice	acquit	
justice	appeal	
justice	arrestJail	yes
justice	chargeIndict	
justice	convict	
justice	execute	
justice	extradite	
justice	fine	
justice	pardon	

¹ Event mention scope is defined as the span of the document from the first event trigger to the next trigger for that same event.

justice	releaseParole	
justice	sentence	
justice	sue	
justice	trialHearing	
life	beBorn	
life	die	yes
life	divorce	
life	injure	yes
life	marry	
manufacture	artifact	yes
movement	transportartifact	yes
movement	transportperson	yes
personnel	elect	yes
personnel	endPosition	yes
personnel	nominate	
personnel	startPosition	yes
transaction	transaction	yes
transaction	transferMoney	yes
transaction	transferOwnership	yes

Table 3. Rich ERE Event Types and Subtypes

Events are assigned a realis attribute of ACTUAL, GENERIC or OTHER. Event arguments are also marked for their realis attributes.

We also label participants for each event mention. Participants are the (previously-labeled) entities and argument fillers involved in that event. We only annotate as participants those entities that are mentioned explicitly within the event mention scope. Argument fillers can come from anywhere in the document. Appendix 2 shows allowable arguments for each event type and subtype. Rich

ERE, however, allows the annotation of event mention triggers even when there are no arguments or participants of the event present in the text.

The same event trigger may be tagged more than once for different event types/subtypes when the trigger instantiate different events or same event type/subtype in certain conjunction cases. This is to allow obligatory inferred events that are in the ERE event taxonomy to be tagged. For example, the trigger “murder” in the example below is the trigger for two Life-Die events, one with the victim “George Besse” and the other with “Rene Audran” and two Conflict-Attack events, one with time argument as 1986 and one with time argument as 1987.

- Cipriani was sentenced to life in prison for the **murder** of Renault chief George Besse in 1986 and the head of government arms sales Rene Audran a year earlier

Tagged event mentions that refer to the same event occurrence will be grouped into Event Hoppers. **Event Hopper** is a more inclusive, less strict notion of event coreference as compared strict event coreference in ACE and Light ERE. Event hoppers contain mentions of events that “feel” coreferential to the annotator even if they do not meet the earlier strict event identity requirement. More specifically, event mentions that have the following features go into the same hopper:

- They have the same event type and subtype (exceptions to this are Contact.Contact and Transaction.Transaction mentions, which can be added to any Contact or Transaction hopper, respectively)
- They have the same temporal and location scope, though not necessarily the same temporal expression or specifically the same date (*Attack in Baghdad on Thursday* vs. *Bombing in the Green Zone last week*)
- Trigger granularity can be different (*assaulting 32 people* vs. *wielded a knife*)
- Event arguments may be non-coreferential or conflicting (*18 killed* vs. *dozens killed*)
- Realis status may be different (*will travel [OTHER] to Europe next week* vs. *is on a 5-day trip [ACTUAL]*)

6 Rich ERE Format

Annotation files are released in an XML format. The entity_mention, relation_mention, and event_mention XML elements all have attributes or contain sub-elements which use character offsets to identify text extents in the source. The offset gives the start character of the text extent and the length attribute is added to the offset to find the string end character. Offset counting starts from the initial character, character 0, of the source document and includes newlines as well as all characters comprising XML-like tags in the source data.

Refer to the DTD for more information.

7 Appendix 1: Allowable arguments for relation types and subtypes

type	subtype	ARG1 Name	ARG1	ARG2 Name	ARG2
physical	locatednear	entity	PER, GPE, LOC, FAC	loc	GPE, LOC, FAC
	resident	per	PER	loc	GPE, LOC, FAC
	orgheadquarter	org	ORG	loc	GPE, LOC, FAC
	orglocationorigin	org	ORG	loc	GPE, LOC, FAC
partwhole	subsidiary	suborg	ORG	parent	GPE, ORG
	membership	member	GPE, ORG	org	GPE, ORG
personalsocial	business	per	PER	per	PER
	family	per	PER	per	PER
	unspecified	per	PER	per	PER
	role	per	PER	role	title
orgaffiliation	employment membership	employeemember	PER	org	GPE, ORG
	leadership	leader	PER	entity	GPE, ORG
	investorshareholder	investorshareholder	PER, ORG, GPE	org	ORG
	studentalum	studentalumni	PER	org	ORG
	ownership	owner	PER	org	ORG
	founder	founder	PER	org	ORG
generalaffiliation	more	per	PER	entity	PER, GPE, LOC

	personage	per	PER	age	age
	orgwebsite	org	ORG	url	url
	opra	org	ORG	entity	PER, ORG

8 Appendix 2: Allowable arguments for event types and subtypes

2016 Evaluation Event Types and Subtypes			
Type	Subtype	ARG_type	allowable ARG entity/filler Type
conflict	attack	attacker	PER, ORG, GPE
conflict	attack	target	PER, GPE, ORG, vehicle, FAC, weapon, commodity
conflict	attack	instrument	weapon, vehicle, commodity
conflict	attack	place	GPE, LOC, FAC
conflict	attack	time	time
conflict	demonstrate	entity	PER, ORG
conflict	demonstrate	place	GPE, LOC, FAC
conflict	demonstrate	time	time
contact	meet	entity	PER, ORG, GPE
contact	meet	entity	PER, ORG, GPE
contact	meet	place	GPE, LOC, FAC
contact	meet	time	time
contact	correspondence	entity	PER, ORG, GPE
contact	correspondence	entity	PER, ORG, GPE
contact	correspondence	place	GPE, LOC, FAC
contact	correspondence	time	time
contact	contact	entity	PER, ORG, GPE
contact	contact	entity	PER, ORG, GPE
contact	contact	place	GPE, LOC, FAC
contact	contact	time	time
contact	broadcast	entity	PER, ORG, GPE
contact	broadcast	audience	PER, ORG, GPE
contact	broadcast	place	GPE, LOC, FAC
contact	broadcast	time	time
manufacture	artifact	agent	PER, ORG, GPE
manufacture	artifact	artifact	vehicle, weapon, FAC, commodity
manufacture	artifact	instrument	weapon, vehicle, commodity
manufacture	artifact	place	GPE, LOC, FAC
manufacture	artifact	time	time
life	injure	agent	PER, ORG, GPE
life	injure	victim	PER
life	injure	instrument	weapon, vehicle, commodity
life	injure	place	GPE, LOC, FAC
life	injure	time	time
life	die	agent	PER, ORG, GPE
life	die	victim	PER
life	die	instrument	weapon, vehicle, commodity

life	die	place	GPE, LOC, FAC
life	die	time	time
movement	transportperson	agent	PER, ORG, GPE
movement	transportperson	person	PER
movement	transportperson	instrument	vehicle, weapon
movement	transportperson	origin	GPE, LOC, FAC
movement	transportperson	destination	GPE, LOC, FAC
movement	transportperson	time	time
movement	transportartifact	agent	PER, ORG, GPE
movement	transportartifact	artifact	weapon, vehicle, commodity, FAC
movement	transportartifact	instrument	vehicle, weapon
movement	transportartifact	origin	GPE, LOC, FAC
movement	transportartifact	destination	GPE, LOC, FAC
movement	transportartifact	time	time
personnel	startposition	entity	ORG, GPE
personnel	startposition	person	PER
personnel	startposition	position	title
personnel	startposition	place	GPE, LOC, FAC
personnel	startposition	time	time
personnel	endposition	entity	ORG, GPE
personnel	endposition	person	PER
personnel	endposition	position	title
personnel	endposition	place	GPE, LOC, FAC
personnel	endposition	time	time
personnel	elect	agent	PER, ORG, GPE
personnel	elect	person	PER
personnel	elect	position	title
personnel	elect	place	GPE, LOC, FAC
personnel	elect	time	time
transaction	transferownership	giver	PER, ORG, GPE
transaction	transferownership	recipient	PER, ORG, GPE
transaction	transferownership	beneficiary	PER, ORG, GPE
transaction	transferownership	thing	vehicle, weapon, FAC, ORG,commodity
transaction	transferownership	place	GPE, LOC, FAC
transaction	transferownership	time	time
transaction	transfermoney	giver	PER, ORG, GPE
transaction	transfermoney	recipient	PER, ORG, GPE
transaction	transfermoney	beneficiary	PER, ORG, GPE
transaction	transfermoney	money	money
transaction	transfermoney	place	GPE, LOC, FAC
transaction	transfermoney	time	time
transaction	transaction	giver	PER, ORG, GPE

transaction	transaction	recipient	PER, ORG, GPE
transaction	transaction	beneficiary	PER, ORG, GPE
transaction	transaction	place	GPE, LOC, FAC
transaction	transaction	time	time
justice	arrestjail	agent	PER, ORG, GPE
justice	arrestjail	person	PER
justice	arrestjail	crime	crime
justice	arrestjail	place	GPE, LOC, FAC
justice	arrestjail	time	time
Event types and Subtypes Not in 2016 Evaluation			
business	startorg	agent	PER, ORG, GPE
business	startorg	org	ORG
business	startorg	place	GPE, LOC, FAC
business	startorg	time	time
business	endorg	org	ORG
business	endorg	place	GPE, LOC, FAC
business	endorg	time	time
business	declarebankruptcy	org	ORG, PER, GPE
business	declarebankruptcy	place	GPE, LOC, FAC
business	declarebankruptcy	time	time
business	mergeorg	org	ORG
business	mergeorg	org	ORG
business	mergeorg	place	GPE, LOC, FAC
business	mergeorg	time	time
life	beborn	person	PER
life	beborn	place	GPE, LOC, FAC
life	beborn	time	time
life	marry	person	PER
life	marry	person	PER
life	marry	place	GPE, LOC, FAC
life	marry	time	time
life	divorce	person	PER
life	divorce	person	PER
life	divorce	place	GPE, LOC, FAC
life	divorce	time	time
personnel	nominate	agent	PER, ORG, GPE
personnel	nominate	person	PER
personnel	nominate	position	title
personnel	nominate	place	GPE, LOC, FAC
personnel	nominate	time	time
justice	releaseparole	agent	PER, ORG, GPE
justice	releaseparole	person	PER

justice	releaseparole	crime	crime
justice	releaseparole	place	GPE, LOC, FAC
justice	releaseparole	time	time
justice	trialhearing	prosecutor	PER, ORG, GPE
justice	trialhearing	adjudicator	PER, ORG, GPE
justice	trialhearing	defendant	PER, ORG, GPE
justice	trialhearing	crime	crime
justice	trialhearing	place	GPE, LOC, FAC
justice	trialhearing	time	time
justice	sentence	adjudicator	PER, ORG, GPE
justice	sentence	defendant	PER, ORG, GPE
justice	sentence	sentence	sentence
justice	sentence	crime	crime
justice	sentence	place	GPE, LOC, FAC
justice	sentence	time	time
justice	fine	adjudicator	PER, ORG, GPE
justice	fine	entity	PER, ORG, GPE
justice	fine	money	money
justice	fine	crime	crime
justice	fine	place	GPE, LOC, FAC
justice	fine	time	time
justice	chargeindict	prosecutor	PER, ORG, GPE
justice	chargeindict	adjudicator	PER, ORG, GPE
justice	chargeindict	defendant	PER, ORG, GPE
justice	chargeindict	crime	crime
justice	chargeindict	place	GPE, LOC, FAC
justice	chargeindict	time	time
justice	sue	plaintiff	PER, ORG, GPE
justice	sue	adjudicator	PER, ORG, GPE
justice	sue	defendant	PER, ORG, GPE
justice	sue	crime	crime
justice	sue	place	GPE, LOC, FAC
justice	sue	time	time
justice	extradite	agent	PER, ORG, GPE
justice	extradite	person	PER
justice	extradite	crime	crime
justice	extradite	origin	GPE, LOC, FAC
justice	extradite	destination	GPE, LOC, FAC
justice	extradite	time	time
justice	acquit	adjudicator	PER, ORG, GPE
justice	acquit	defendant	PER, ORG, GPE
justice	acquit	crime	crime

justice	acquit	place	GPE, LOC, FAC
justice	acquit	time	time
justice	convict	adjudicator	PER, ORG, GPE
justice	convict	defendant	PER, ORG, GPE
justice	convict	crime	crime
justice	convict	place	GPE, LOC, FAC
justice	convict	time	time
justice	appeal	prosecutor	PER, ORG, GPE
justice	appeal	adjudicator	PER, ORG, GPE
justice	appeal	defendant	PER, ORG, GPE
justice	appeal	crime	crime
justice	appeal	place	GPE, LOC, FAC
justice	appeal	time	time
justice	execute	agent	PER, ORG, GPE
justice	execute	person	PER
justice	execute	crime	crime
justice	execute	place	GPE, LOC, FAC
justice	execute	time	time
justice	pardon	adjudicator	PER, ORG, GPE
justice	pardon	defendant	PER, ORG, GPE
justice	pardon	crime	crime
justice	pardon	place	GPE, LOC, FAC
justice	pardon	time	time