DEFT Rich ERE Annotation Guidelines:

Argument Filler V2.1

Linguistic Data Consortium
December 8, 2014

1. Basic Concepts

The purpose of this annotation project is to mark up texts for entities, coreference, events and relations. The primary purpose is for the annotations to describe the meaning of the text, as opposed to its syntactic or lexical aspects. The annotation is carried out level by level. This document describes the ad hoc annotation of Argument Filler during Relation and and Event annotation tasks.

Argument Filler annotation is not a standalone annotation task like Entity, Relation or Event, but rather annotation during Relation and Event tasks to detect and characterize arguments for certain Relations and Events that haven't been captured in the Entity annotation task.

1.1 What is an Argument Filler?

An Argument Filler is a string of text that can fill in as arguments for certain Relations or Events. They are only taggable when they are arguments of designated Relations or Events.

We will only be interested in certain types of possible Argument Fillers. Specifically, we will be annotating AGE, MONEY, URL, CRIME, SENTENCE, WEAPON, VEHICLE, COMMODITY, TITLE and TIME. Details of each Argument Filler type will be discussed in Section 2.

1.2 Taggability

It is important to note that there are two types of arguments that we are calling Argument Fillers for this task:

1. Strings that participate as arguments in Relations. These strings are only taggable when they are arguments of certain Relations.

Relation Argument Fillers: Only tagged when used as an argument in a Relation		
Filler Type	Argument Role	Relation Type and Subtype
AGE (AGE)	age	General-Affliation.Person-age
URL (URL)	url	General-Affliation.Org-website
TITLE (TTL)	role	Personal-Social.Role

2. Strings that participate as arguments in Events. These strings are only tagged when they are arguments of a taggable Event.

Note: TITLE can either be Relation Argument Filler or Event Argument Filler.

Event Argument Fillers: Only tagged when used as an argument in an Event		
Filler Type	Argument Role	Event types and subtypes
MONEY (MON)	money	Movement Transaction Justice.Fine
WEAPON (WEA) VEHICLE (VEH) COOMODITY (COM)	target instrument	Conflict.Attack
	artifact	Manufacture.Artifact
	instrument	Movement.TransportPerson
	artifact instrument	Movement.TransportArtifact
	instrument	Life.Injure
	instrument	Life.Die
	thing	Transaction.TransferOwnship
TITLE (TTL)	title	all Personnel type events
SENTENCE (SEN)	sentence	Justice.Sentence
CRIME (CRM)	crime	all Justice type events
TIME (TME)	time	all event types.subtypes

1.3 Extent

The rules for identifying the extent of an Argument Filler will vary from type to type. The specific extent rules for a given type will be provided in the section describing that type. However, some general properties of all Argument Filler extents can be identified.

Many Argument Fillers are mentioned by a noun phrase (NP). The Entity task guidelines introduce a detailed account of the manner in which the full extent of an NP can be identified. That account is repeated here for convenience.

The extent of an Argument Filler consists of the entire NP unless otherwise specified. In cases where it's difficult to decide if modifiers should be included or not, the Argument Filler extent should be maximally inclusive. In the case of a discontinuous constituent, the extent goes to the end of the constituent, even if that means including tokens that are not part of the constituent. Thus, in:

The terrorist was charged with conspiracy yesterday in the bombing of the USS Cole.

the extent of a "crime" filler for Justice.Charge-Indict event is:

[conspiracy yesterday in the bombing of the USS Cole]$_{CRM}$

The extent of an "artifact" filler of Conflict.Attack event is a VEHICLE type:

[the USS Cole]$_{VEH}$

The extent includes all the modifiers of a NP, including determiners, prepositional phrases and relative clauses.

Generally speaking, tokens are broken at white space, and each item of punctuation is treated as a separate character. As a rule, we do not include punctuation such as commas, periods, and quotation marks in the extent of a mention unless words included within the extent continue on after the punctuation mark. Possessive endings ('s) are treated as separate tokens, and contractions are split (so that "*we're*" becomes the two tokens "we" and "*'re*"). Extents must begin at the beginning of a token and end at the end of a token.

Unlike Entities, Argument Fillers do not have heads, mention type (NOM, NAM, PRO) or mention class (SPC, NonSPC).

2. Argument Filler Types

Note: In the examples of this section, [square brackets] are used to indicate the extent of [Argument Fillers] and **bold** font is used to indicate **Event triggers**. The following examples will only include Events and Event triggers when an Argument Filler corresponds to the Event..

2.1 AGE (AGE)

AGE is an argument filler for General-Affliation.Person-age. It is limited to a PERSON entity's age. The extent of AGE only needs to include the string of texts that describe the age.

> [a [10-year-old]_{AGE} girl]_{PER} was the only victims
>> General-Affliation.Person-age
>> ARG1: a 10-year-old girl; ARG2: 10-year-old

> McCain, at [the age of 74]_{AGE}, became the oldest nominee
>> General-Affliation.Person-age
>> ARG1: McCain; ARG2: the age of 74

2.2 URL (URL)

A URL is an argument filler for General-Affliation.Org-website. It is limited to the official URL of an ORGANIZATION entity. A URL type is taggable whenever the virtual location of a webpage of an ORG type Entity is provided.

The extent of an URL mention is the smallest sequence of tokens such that all of the tokens in the string are included. URL Argument Fillers should be contiguous but will contain tokens of various types (e.g. symbols, numbers and letters). Some times the URL is a full path of the resource. In such case, only the top level URLs are tagged as the URL.

> [http://www.upenn.edu]/computing/_{URL}

> [google.com]_{URL}/voice

> [http://www.kungfunecktie.com]_{URL}

NOTE: Former websites are acceptable URLs. Also, websites may be extracted from email addresses that are clearly tied to the organization, such as [cnn.com] from the email address, "FareedZakaria@cnn.com".

2.3 TITLE (TTL)

A TITLE is taggable when it is the "role" filler of Personal-Social.Role relation or "title" filler of any Personnel type events.

Titles are the personal titles and honorifics, official rank or status, and specific employed occupations or professional positions. It is used to establish a particular relationship between a PER mention and the title, honorific, position, or occupation, both in the same sentence. Annotators will need to use best judgment when determining whether a position or occupation is specific enough to be tagged as a

title (e.g., general types such as worker, official, member, employee, will not be tagged TTL).

For our purposes, the extent of a TTL is limited to the string of text that describes the position or rank itself, <u>independent of</u> its organizational circumstances or relationships, and <u>excluding</u> a preceding definite article ('the') and any other pre-posed or post-posed modifiers.

Because TTLs occur in conjunction with other entities and refer to other entities, they have some special rules. Title elements are often but not always adjacent to a PER entity, which can be of mention level NAM, NOM or PRO. The most frequent constructions TTLs occur in are title + name, appositives, and copula constructions.

Just as with appositives, all titles, positions, and honorifics will be marked <u>separately</u> from the individual's name. For instance, in the following sentences, there are two separate entities marked:

- [Vice President]$_{TTL}$ [Biden]$_{PER}$
        Personal-Social.Role relation
        ARG1: Biden; ARG2: Vice President

- [spokesperson]$_{TTL}$ [Mary Gillette] $_{PER}$
        Personal-Social.Role relation
        ARG1: Mary Billette; ARG2: spokesperson

- [Michelle Obama] $_{PER}$, [First Lady] $_{TTL}$, spoke at the event.
- [Michelle Obama] $_{PER}$, [who]$_{PER}$ is the [First Lady] $_{TTL}$, spoke at the event.
- [She]$_{PER}$ is the [First Lady] $_{TTL}$.
        Personal-Social.Role relation
        ARG1: Michelle Obama; ARG2: First Lady

- The famous Greek [philosopher]$_{TTL}$ [Zeno]$_{PER}$ was born in 490 BC.
        Personal-Social.Role relation
        ARG1: Zeno; ARG2: philosopher

Sometimes the name of the person is split into two pieces by the title. In these cases, we will annotate a nested title.

- [Alfred [Lord]$_{TTL}$ Tennyson]$_{PER}$
        Personal-Social.Role relation
        ARG1: Alfred Lord Tennyson; ARG2: Lord

TTLs are often mentioned in tandem with the organization (ORG) or geo-political (GPE) entity to which they pertain, e.g.: "Mayor of Boston", "Vice President of Marketing", "the Dover Rotary Club's treasurer", "the commissioner of the National

Basketball Association", etc. Thus, if the phrase containing the position title itself also contains another taggable entity type, be sure to tag each entity separately (TTL-ORG relationships will be tagged in the Relations level of annotation). Examples:

- [Treasury]$_{ORG}$ [Secretary]$_{TTL}$ [Jackson]$_{PER}$
- [Justice]$_{ORG}$ [Minister]$_{TTL}$ [Giovannia Maria Flick]$_{PER}$
- [Mission Control]$_{ORG}$ [Chief]$_{TTL}$ [Vladimir Solovyov]$_{PER}$
- [[US]$_{GPE}$ Army]$_{ORG}$ [negotiator]$_{TTL}$ [Harold Norman]$_{PER}$
- [US]$_{GPE}$ [Secretary]$_{TTL}$ of [State]$_{ORG}$ [Hillary Clinton]$_{PER}$

Compound and multi-word titles may be tagged so that their mention extents include all words essential to the understanding of the title or position itself, excluding an adjacent organization or GPE mention, e.g.:

- As the newly **appointed**$_{Personnel.StartPosition}$ [Assistant Deputy Sheriff]$_{TTL}$, it is [Bob]$_{PER}$'s duty to inquire into certain deaths.
  Personal-Social.Role relation
  ARG1: Bob; ARG2: Assistant Deputy Sheriff

- Japan's parliament **elected** $_{Personnel.Elect}$ [him]$_{PER}$ [prime minister]$_{TTL}$.
  Personal-Social.Role relation
  ARG1: him; ARG2: prime minister

- White House [Chief of Staff]$_{TTL}$ [Jack Lew]$_{PER}$ **resigned** $_{Personnel.EndPosition}$ last year.
  Personal-Social.Role relation
  ARG1: Jack Lew; ARG2: Chief of Staff

- **former** $_{Personnel.EndPosition}$ [President]$_{TTL}$ Bill Clinton
  Personal-Social.Role relation
  ARG1: Bill Clinton; ARG2: President

The only exception to excluding an ORG or GPE mention from a title mention extent is when one is nested within it:

- [Deputy [Foreign]$_{ORG}$ Minister]$_{TTL}$

In some cases, more than one title will be presented for a single person. In case all of these titles are adjacent to the person name, we will tag each of the titles separately as TTL:

- [Karachi]$_{GPE}$ [Mayor]$_{TTL}$ and [GlobalCom]$_{ORG}$ [Chairman]$_{TTL}$ [Mr]$_{TTL}$ [Anwar Ayub]$_{PER}$
  Personal-Social.Role relation
  ARG1: Anwar Ayub; ARG2: Mayor

Personal-Social.Role relation
ARG1: Anwar Ayub; ARG2: Chairman

Personal-Social.Role relation
ARG1: Anwar Ayub; ARG2: Mr

In other cases, the same tiltle may be associated for different persons. Number agreement is not essential and a plural title may be associated with a singular PER where necessary.

- [Leaders]TTL [Ayrault]PER and [Merkel]PER met in Berlin.
    Personal-Social.Role relation
    ARG1: Ayrault; ARG2: Leaders

    Personal-Social.Role relation
    ARG1: Merkel; ARG2: Leaders

NOTE: When a string that may serve as a title or position is being used as a nominal to refer to a person who is not named in relation to it within the sentence, the entity mention should be <u>tagged as PER and not as TTL</u>. For instance, consider these two contrasting cases:

- [Joe Biden]PER is [Vice President]TTL
- [The strongest supporter]PER was [the vice president]PER.

NOTE: Phrases that establish ad hoc, informal, or general roles—as opposed to more specific, formal positions, occupation, titles, and honorifics—should <u>not</u> be tagged as titles:

- Suzie, the ~~[brain]~~TTL in [the family]PER., explained how the machine worked.
- ~~[Class clown]~~TTL John made a joke.
- Amy was always the ~~[philosopher]~~TTL in our discussions.

2.4 MONEY (MON)

A MONEY Argument Filler is tagged whenever it is associated with Transaction and Justice.Fine events. It is described in terms of the currency of some country or region (e.g. *US Dollars* or *Euros*). The extent of a Money mention needs to include modifying strings as well as units.

*[nearly $400 Million] was **given**Transaction.Transfer-Money to the charity in 2010*

*He was **fined**Justice.Fine [50 dollars]for turning right on red*

*[20 Euros] **bought**Transaction.Transfer-Money him a nice meal in Granada!*

2.5 CRIME (CRM)

A CRIME Argument Filler will be tagged whenever the offense associated with some JUSTICE Event is explicitly expressed. There must be a taggable instance of the corresponding JUSTICE Event for there to be a taggable CRIME Argument Filler.

Since most CRIME Argument Fillers will be expressed in the form of an NP, the extent of CRIME will be defined generally, as in Section 1.2 above.

- The terrorist was **charged** $_{Justice.ChargeIndict}$ with [conspiracy in the bombing of the USS Cole]$_{CRM}$.

- 46-year-old Abu Talib was **sentenced** $_{Justice.Sentence}$ to life imprisonment in 1990 in Sweden for [terrorist acts in Amsterdam, Copenhagen and Stockholm between 1985 and 1986]$_{CRM}$.

2.6 Sentence (SEN)

A SENTENCE Argument Filler will be mentioned whenever a sentence for the JUSTICE.Sentence Event is explicitly expressed. There must be a taggable instance of the corresponding Justice.Sentence Event for there to be a taggable SENTENCE Argument Filler.

Since most SENTENCE Argument Fillers will be expressed in the form of an NP, the extent of SENTENCE will be defined generally, as in Section 1.2 above.

- 46-year-old Abu Talib was **sentenced** $_{Justice.Sentence}$ to [life imprisonment]$_{SEN}$ in 1990 in Sweden for terrorist acts in Amsterdam, Copenhagen and Stockholm between 1985 and 1986.

- Solomon could be **sentenced** $_{Justice.Sentence}$ to [up to 211 years in prison]$_{SEN}$.

- Hutomo ``Tommy'' Mandala Putra, 37, was **sentenced** $_{Justice.Sentence}$ to [18 months in prison] $_{SEN}$ on Sept. 22 by the Supreme Court, which overturned an earlier acquittal by a lower court.

2.6 Vehicle (VEH)

A VEHICLE is a physical device primarily designed to move an object from one location to another, by (for example) carrying, pulling, or pushing the transported object. Vehicle entities may or may not have their own power source.

Since most VEHICLE Argument Fillers will be expressed in the form of an NP, the extent of VEHICLE will be defined generally, as in Section 1.2 above.

A VEHICLE Argument Filler will be mentioned whenever a vehicle that is associated with an Argument Filler of the following event types and subtypes is explicitly expressed:

        Conflict.Attack: Target or Instrument
        Manufacture.Artifact: Artifact
        Life.Injure: Instrument
        Life.Die: Instrument
        Movement.TransportPerson: Instrument
        Movement.TransportArtifact: Artifact or Instrument

There must be a taggable instance of a Relation or Event mentioned above for there to be a taggable Vehicle Argument Filler. Compare the following examples:

- The very next day, he **boarded** $_{\text{Movement.TransporPerson}}$ [a passenger train bound for Amristar].
- What if [The Chinatown bus] showed up late, so everyone was pretty ticked off.

The first example does not include any Relation or Event mention, whereas the second contains a Movement.TransportPerson event. Below are more examples:

- He used [that really broken down bike] to **get** $_{\text{Movement.TransportPerson}}$ to school.

- the Oct. 12th **bombing** $_{\text{Conflict.Attack}}$ of the [U.S.S. Cole].

- She **rode** $_{\text{Movement.TransportPerson}}$ [a jet ski] from the pier to the shore
- Most passengers **gathered** $_{\text{Movement.TransportPerson}}$ in the [the food compartment of [the train]].
- The Army **purchased** $_{\text{Transaction.TransferOwnership}}$ another [202,000 vehicles].
- Troops **arrived** $_{\text{Movement.TransportPerson}}$ in town via [Israeli tanks and gunships]

2.7 Weapon (WEA)

A WEAPON is a physical device that is primarily used as an instrument for physically harming or destroying entities.

Since most WEAPON Argument Fillers will be expressed in the form of an NP, the extent of WEAPON will be defined generally, as in Section 1.2 above.

A WEAPON Argument Filler will be mentioned whenever a weapon that is associated with an Argument Filler of the following event types and subtypes:

Conflict.Attack: Target or Instrument
Manufacture.Artifact: Artifact
Life.Injure: Instrument
Life.Die: Instrument
Movement.TransportPerson: Instrument
Movement.TransportArtifact: Artifact or Instrument

There must be a taggable instance of a Relation or Event mentioned above for there to be a taggable Weapon Argument Filler.

Compare the following examples:
- [AK-47] is the best-selling assault weapon.
- A gang of 5 masked man **stormed** Conflict.Attack the bank with [their AK-47s]...

The first example does not include a Relation or Event mention, whereas the second contains a Conflict.Attack event as well as an Agent-Artifact.UOIM relation. Below are more examples:

- He **robbed** *Transaction.TransferOwnership* a bank, holding up the entire place with [a gun].
- [Gunfire] shelled the shore on D-Day.
  - annotate as the trigger of Conflict.Attack event, not a Weapon entity
- [Tanks] **blew up** Conflict.Attack and crushed a café in the Kasbah.
- He was **beaten** Conflict.Attack repeatedly with [a baseball bat]

- An international investigation led to charges against two alleged Libyan intelligence agents of **putting** Conflict.Attack [a bomb] on the plane.

- [Sarin gas] **filled** Conflict.Attack the room without any warning

- Police **fired** Conflict.Attack [tear gas] at the unarmed protestors

NOTE: VEHICLES that are manufactured to also be weapons (like tanks, bomber aircrafts, etc.) will be tagged as WEAPON Argument Fillers, but normal locomotive vehicles that used such that they harm people will still be tagged as VEHICLES (as in "he **hit**Conflict.Attack me with [his car]VEH).

2.8 COMMODITY (COM)
A commodity is a product or article of trade for which someone pays or barters.

Since most COMMODITY Argument Fillers will be expressed in the form of an NP, the extent of COMMODITY will be defined generally, as in Section 1.2 above.

A Commodity Argument Filler will be mentioned whenever a commodity that is associated with an Argument Filler of the following event types and subtypes:

Conflict.Attack: Instrument
Manufacture.Artifact: Artifact, Instrument
Life.Injure: Instrument
Life.Die: Instrument
Movement.TransportArtifact: Artifact, Instrument
Movement.TransportPerson: Artifact, Instrument

There must be a taggable instance of a Relation or Event mentioned above for there to be a taggable Commodity Argument Filler.

Compare the following examples:

- There's [a bunch of bananas] sitting on the curb outside.
- I **bought** [a bunch of bananas]COM at the store this morning.

The first example does not include an Event mention, whereas the second contains a Transaction.Transfer-Ownership event. Below are more examples:

- [Several storage containers of firearms and ammunition] were **confiscated** Transaction.TransferOwnership.
- He was found with [ammonium nitrate] inside his van. (No Event).
- [Five kilos of heroin] were located and **removed**Movement.TransportArtifact from the perp's home.
- [A large amount of fertilizer] in his backyard was **moved**Movement.TransportArtifact to his basement where he began creating explosives.

2.9 TIME (TME)

A Time is a temporal expression that can reference a *calendar date*). Only temporal expressions that are or can be resolved to calendar dates should be tagged as TIME Argument Fillers. The extent of TIME only needs to include the string of texts that describe the time.

There must be a taggable instance of a taggable Event for there to be a taggable Time Argument Filler. A Time Argument Filler will be tagged whenever a temporal expression that can be placed on a timeline is associated with a taggable Event.

Compare the following examples:

- Hurricane Sandy landed in New Jersey [October 29]
- The **bombing**Conflict.Attack took place on [October 29].

The first example does not include an Event mention, whereas the second contains a Conflict.Attack event.

Taggable temporal expressions include things like 'yesterday', 'Tuesday', etc. – as long as the document publication date is available and can be used to infer the calendar date referred to by the temporal expression. We will both tag the text string referring to the date and normalize it into a standard YYYY-MM-DD format. Below are more examples:

- She **moved** to California in <u>August 2003</u>. = 2003-08-XX
- Disney **purchased** Pixar in <u>2006</u>. = 2006-XX-XX
- The leaders **met** <u>last week</u> in Lagos. (article date: May 25, 2012, "last week" being May 13-19) = 2012-05-1X
- The <u>24 to 26 October 2000</u> **attacks** left 203 **dead**... = 2000-10-2X
- Riots **occurred** <u>between December 10th and the 15th</u> (article date: December 20, 2009) = 2009-12-1X
- The company had given **donations** for <u>the past 10 years</u> (article date: June 20, 2011) = 20XX-XX-XX

- Congress **convenes** [next Thursday].

- Police said the 31-year-old Britain **died** [Thursday].

- They **had lunch** together [two weeks ago today].

- I'm **leaving** on vacation [two weeks from next Tuesday].