Definition of the Annotation Module

Deliverable D3.1

Version FINAL

Authors: Sara Tonelli[1], Rachele Sprugnoli[1]

Affiliation: (1) FBK

NewsReader

POST HOC ERGO PROPTER HOC

COOPERATION

Building structured event indexes of large volumes of financial and economic data for decision making

ICT 316404

Grant Agreement No.	316404
Project Acronym	NEWSREADER
Project Full Title	Building structured event indexes of large volumes of financial and economic data for decision making.
Funding Scheme	FP7-ICT-2011-8
Project Website	http://www.newsreader-project.eu/
Project Coordinator	Prof. dr. Piek T.J.M. Vossen VU University Amsterdam Tel. + 31 (0) 20 5986466 Fax. + 31 (0) 20 5986500 Email: piek.vossen@vu.nl
Document Number	Deliverable D3.1
Status & Version	FINAL
Contractual Date of Delivery	June 2013
Actual Date of Delivery	July 2013
Type	Prototype
Security (distribution level)	Public
Number of Pages	48
WP Contributing to the Deliverable	WP3
WP Responsible	FBK
EC Project Officer	Susan Fraser
Authors: Sara Tonelli[1], Rachele Sprugnoli[1]	
Keywords: Annotation guidelines, annotation tools, events and entities	
Abstract: In this document, we describe the guidelines proposed for annotating events, participants, relations and temporal information in NEWSREADER data. Annotation is detailed both at 'mention' (i.e. textual) and at 'instance' (i.e. semantic) level, and is designed so as to be largely language independent. We also describe the tools that will be used to perform the two types of annotation mentioned above: CAT (CELCT Annotation Tool) will allow annotators to identify at document level all information needed to capture the described events, while CROMER (CROss-document Main Event and Entity Recognition) allows annotating co-referential events and entities in clusters of documents. The two annotation workflows have been devised in order to be easily interlinked and allow merging the top-down and the bottom-up information in a single representation.	

Table of Revisions

Version	Date	Description and reason	By	Affected sections
0.1	Begin May 2013	Draft of paper structure and introduction	Sara Tonelli	1
0.2	Mid May 2013	Guidelines for annotation at document and at corpus level	Rachele Sprugnoli	3, 4
0.3	End of May 2013	Standards for Entity, Event and Temporal Annotation	Sara Tonelli	2
0.4	Mid June 2013	Description of Annotation Tools	Rachele Sprugnoli	5, 6
0.5	Mid June 2013	Final check and additions before internal review	Sara Tonelli	all
0.6	Begin July 2013	Internal review	Itziar Aldabe, German Rigau	all
0.7	9th July 2013	Integrate reviews in final version, added Executive Summary and Conclusions	Sara Tonelli, Rachele Sprugnoli	all
1.0	date	approval by project manager	who	-

Executive Summary

The present document provides a description of the tools developed to annotate information on events and entities required within NewsReader, as well as a detailed discussion of the annotation choices made to address the project's requirements.

The document consists of three parts: we first provide an overview of annotation standards proposed in the literature that may be relevant to NewsReader annotation. These standards concern the basic *elements* to be modeled in the project's system, i.e. entities, events and temporal expressions. We further focus on the *relations* among these elements, with a particular emphasis on causal and temporal relations between events, which are crucial to building storylines.

Based on the state of the art, we detail our proposal for annotating event-based information both at document and at cluster level. The former annotation is largely inspired by TimeML, with some extensions mainly concerning the distinction between mention (i.e. textual) and instance (i.e. semantic) level of annotation, and the integration of participants' information. We call this annotation process *bottom-up*, abstracting from mentions to instances. In the cluster-based annotation, instead, the direction is reversed, or *top-down*: annotators are first asked to fill in a template that uniquely describes events or entities, linked to an external knowledge source (e.g. DBPedia), and then to select clusters of documents where such event or entity is mentioned.

In order to guarantee enough flexibility and language-independence in our annotation workflow, we extend two existing annotation tools developed at FBK, taking into account NewsReader requirements. For the bottom-up approach, we use CAT (CELCT Annotation Tool) and introduce several specific features, the most important one being the possibility to handle event coreference at document level. For the top-down approach, we enrich the CROMER tool (CRoss-document Main Events and entities Recognition) with project-specific extensions, and the most relevant one is again the module to annotate cross-document coreference for events.

The activities detailed in this document are the outcome of a continuous collaboration between the consortium partners, especially FBK, VUA and EHU. Indeed, annotation choices are extremely relevant to the activities concerning NLP processing and semantic modelling (WP4 and WP5). Besides, the annotation choices have been made in order to take into account recent advances in event annotation (see FBK and VUA participation to the First Workshop on Events), and to benefit from the partners' experience in defining annotation standards in past projects such as KYOTO (FP7-ICT-211423) and TERENCE (FP7-ICT-2010-25410).

Contents

1 Introduction

The goal of the NewsReader project (ICT-316404) is to develop a system to support decision makers in the economic-financial domain by extracting event streams and building event timelines. This has to be performed starting from large volumes of data in the 4 languages of the project, namely Dutch, English, Spanish and Italian. Given the task complexity and the number of natural language processors to be integrated in the system, it is crucial to develop gold data to allow on the one hand the evaluation of such processors, and on the other hand to adapt them to the specific project domain.

In order to support the creation of such gold data, existing standards for temporal and event annotation have been revised and a novel annotation standard has been proposed to fit the project needs. Besides, two annotation tools have been developed to support the required annotation activities. Both tools have been designed so as to be language-independent, and to be accessible through a web-browser.

The main difference between NewsReader annotation and other existing linguistic frameworks is that it aims at developing a model in which linguistic information and the underlying semantic layer are interconnected, possibly including non-linguistic sources. This connection will be used to reason over the cumulated and linked sources of knowledge and information to interpret the often incomplete and fragmented information that is provided by each source.

In order to comply with this model, annotation will be carried out from two different perspectives. Annotation at document level is performed *bottom-up*, and is driven by linguistic evidence. Mentions are marked in the text as well as relations between them (participants, time, temporal and causal relations, intra-document coreference). However, in order to allow interconnection with RDF models and semantic web standards, annotation is also performed in a *top-down* fashion: entities and events are first described through info-boxes similar to Wikipedia ones, directly linked to external knowledge sources (e.g. DBPedia); then they are connected to document clusters where such entities and events are mentioned.

This deliverable introduces the choices that have been taken at a theoretical level to devise the NewsReader annotation standard and account for the event modeling needs. Then, it describes a first version of the two modules to perform annotation at document and at corpus level. These have been inspired by existing linguistic standards for event and temporal annotation (e.g. TimeML, FrameNet, KAF, TAF) as well as by event models proposed within the Semantic Web community (e.g. the Simple Event Model).

At document level, we have developed an extension of *CAT* (CELCT Annotation Tool), through which the annotator can distinguish between the *mention* and the *instance* layer, both for events and for entities. Additional functionalities for computing annotation statistics and inter-annotator agreement have been implemented.

At corpus level, the *CROMER* tool to perform cross-document coreference between named entities has been extended and improved in order to cope with cross-document coreference between events in a top-down fashion. The tool has been designed so as to guarantee again the highest interconnection with external resources and ease event modelling and chaining. Both tools can be accessed through a web browser, but access is currently restricted. In order to get a username and password, please contact satonelli@fbk.eu.

This deliverable is structured as follows. In the next section, we will provide an overview of the linguistic standards for event and temporal annotation which inspired our choices. In Section 3 we present our proposal for annotating events and temporal information at document level. In Section 4 we present our proposal for annotating cross-document coreference between entities and between events. In Section 5 we introduce the CAT tool and detail the adaptations performed for the NewsReader project, while in Section 6 we present the CROMER tool implemented for annotating cross-document coreference. Finally, we draw some conclusions in Section 7.

2 Standards for Entity, Event and Temporal Annotation

Annotation of events and of relations between them has a long tradition in Natural Language Processing (NLP). Therefore, we analyzed several existing guidelines for event, participants and temporal modeling, so that we could take advantage of long-standing standards and tailor them to NewsReader requirements.

2.1 Entities, Events and Temporal Expressions

The basic story elements in news are the *whats*, *whens*, *wheres* and *whos*, since they are necessary to build a storyline of what took place in the news description.

Although the notion of **event**, i.e. *what* happened, may seem intuitive, there is no consensus in the NLP community on which linguistic expressions describe events and should be annotated as such. Most attempts to annotate events have been limited to specific scenarios or domains, as in Linguistic Data Consortium's Automatic Content Extraction (ACE) and Machine Reading initiatives ([Bagga and Baldwin, 1998], [He, 2007]). The recent OntoNotes annotations include more general event mentions, which however are still limited to verbs and nominalizations [Weischedel *et al.*, 2010]. Events are also a crucial element in TimeML framework [Pustejovsky *et al.*, 2010], where they are a cover term for a large number of different elements including tensed or untensed verbs, nominalizations, adjectives, predicative clauses and prepositional phrases. In the light of this uncertain definition of events, a specific workshop[1] has been organized and co-located with NAACL2013 conference, with the goal of analyzing agreement and disagreements on annotation schemes for events and achieving a consensus on the treatment of events. The NewsReader consortium participated to the workshop and to the shared task, proposing the two-layered approach to annotation presented in Sections 3 and 4. Given the specific domain of NewsReader and the large amount of data to be processed, we limit annotation to three main event types: *i)* speech acts and cognitive events, *ii)* contextual events relevant to the financial domain, and *iii)* 'grammatical' events such as *take place*, *occur*, etc. (more details on this in Section 3.2 below).

In addition to events, the interpretation of expressions that refer to *time* is crucial to build a storyline. Such expressions tell us not only when something happened but also for how long something lasted and how often something occurs. The DARPA-sponsored evaluations at the end of the '90 (MUC) specified the first annotation standard for the semantic representation of temporal expressions introducing the <TIMEX> tag. Starting from 2000, a new standard based on a new tag called <TIMEX2> [Ferro *et al.*, 2001] [Ferro *et al.*, 2005] was developed under the TIDES program and adopted in the Temporal Expression Recognition and Normalization task within the ACE initiative in 2004 [Linguistic Data Consortium, 2004b]. The TIMEX2 standard contained many advances with respect to the MUC schema. In particular, it covered a wider range of expressions, in-

[1] `https://sites.google.com/site/cfpwsevents/`

cluding context dependent expressions (such as *next Monday, now* that need an anchoring time to be interpreted) and it introduced the attribute VAL that captures the normalized form of the temporal expression following the ISO 8601 standard format. More recently, the <TIMEX3> tag included in TimeML has extended the previous schemes through a greater number of attributes for a more detailed annotation of temporal expressions. For example, it allows to insert start and end points for durations and intervals and to specify the quantifier and time granularity of a temporal expression. In NewsReader we have decided to adopt a version of TIMEX3 simple enough to be executed and yet precise enough for the purposes of the project: thus only the two most relevant attributes are introduced, that is TYPE for temporal expression classification and VALUE for time normalization (for more details see Section 3.3 below).

Answer about *who* is involved in an event and *where* this event happens means recognizing the entities involved in the event itself and identifying the references to these entities within one or more texts. Early work on entity recognition and classification formulated the task as recognizing only proper names and focused on a limited number of entity types: for example, MUC-7 evaluations considered person, location and organization and, later, CONLL-2003 added a miscellaneous type. The ACE program introduced not only a wider set of 7 types of entities (namely, person, location, organization, geo-political entities, weapons, vehicles and facilities) but also a specific annotation scheme and evaluation task for the recognition of intra-document and cross-document coreference (for details on coreference see the subsection "Relations between entities" below). In particular, an *entity* is defined in ACE as an object or set of objects in the world and a *mention* as a textual reference to an entity: annotators identify all mentions of each entity within a document, whether named, nominal or pronominal, label the head of each mention, tag entities and mentions following specific categorizations and cluster the mentions for each entity together into a unique entity ID. The most recent large scale annotation of entities was performed under the OntoNotes project [Weischedel *et al.*, 2010] that expands the ACE set of entities to be tagged defining 11 classes namely, persons, nationalities or religious or political groups, facilities, organizations, geo-political entities, locations, products, named events, works of art, named documents made into laws, and named languages. In NewsReader we simplify the classification asking annotators to identify only 5 classes of entities (i.e. person, organization, location, artifact, and financial) but, at the same time, we pay particular attention to terms widely used in the financial domain. As for mentions, we have decided to adopt the same syntactic types described in the ACE guidelines (more information are given in Section 3.1).

2.2 Relations between events and entities

Although knowing which participants are involved in an event is crucial to unambiguously identify such event, TimeML [Pustejovsky *et al.*, 2010] does not foresee participants' annotation. Some attempts to add argument information to TimeML have been made in the past [Pustejovsky *et al.*, 2006a; Pustejovsky *et al.*, 2007], but this has not led to the annotation of a large-scale corpus or to the extension of TimeBank [Pustejovsky *et al.*, 2006b] with such information.

Nevertheless, research on semantic roles and predicate arguments has a long-standing tradition in linguistics and in NLP, though it has been conducted in parallel to research on events and temporal processing. Relations between a predicate and its arguments can be assessed primarily at syntactic level, using different parsing formalisms. In the *Penn TreeBank* [Marcus *et al.*, 1994], for instance, constituency-based annotation of sentences included some basic predicate argument structure, including the subjects, times and locations of verbal events signaled by "functional tags" of syntactic constituents (e.g. -SBJ, -TMP, -LOC). When a syntactic dependency formalism is applied, rather than syntactic constituency, predicate argument structure is often visible directly in the dependencies. For instance, the Stanford parser[2], based on *typed dependencies* [de Marneffe and Manning, 2008b], would produce the syntactic analysis displayed in Fig. 1 for the sentence "The Council will undergo extensive turnover this year with some 21 of its 51 members stepping down". The main event "undergo" is connected to the nominal subject "Council", the direct object "turnover" and two other modifiers, one of which is marked as temporal (*tmod*). The advantage of this representation compared to the constituency-based one is that events are connected directly to the heads of the participants, making it easier to acquire information on the semantics of the argument fillers. Besides, relations between events and participants are labeled, marking the distinction between mandatory arguments (subjects, objects, etc.) and modifiers. This information is very relevant to the NewsReader system, because they contribute, for instance, to identifying event chains and event coreference.

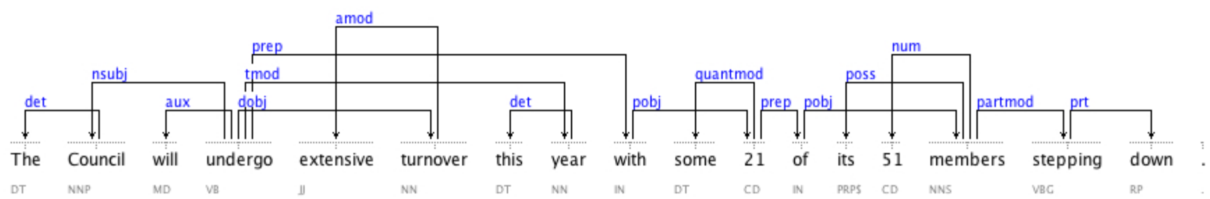

Figure 1: Example sentence with typed dependencies

Other annotation initiatives have focused more on semantic roles, enriching syntactic structure with information on the role played by the arguments e.g. Agent, Patient, Instrument, etc. The most notable example is the PropBank corpus [Palmer *et al.*, 2005], where verb structures have been manually annotated within a sentence context and enriched with argument labels (Arg0, Arg1, Arg2, etc.) and functional tags attached to the verb modifiers (MNR, LOC, TMP and others). The goal of the initiative was to provide consistent argument labels across different syntactic realizations of the same verb. Fig 2 displays the output of *Mate* semantic role labeler[3] trained on PropBank on the same example sentence.

Another framework to model semantic roles has been proposed within the FrameNet

[2] http://nlp.stanford.edu/software/lex-parser.shtml
[3] http://barbar.cs.lth.se:8081/

	The	Council	will	undergo	extensive	turnover	this	year	with	some	21	of	its	51	members	stepping	down
undergo.01	A0		AM-MOD		A2		AM-TMP		A2								
turnover.02					A2				A2								
member.01													A1				
step.01										A1							AM-DIR

Figure 2: Example sentence with semantic roles from PropBank

project [Baker *et al.*, 2003], in which predicates are first grouped according to semantic similarity and then for each of these groups, called *frames*, a set of frame-specific roles has been defined. For instance, in the example above, "undergo" would be part of the UNDERGOING frame. Its arguments would be labeled as follows:

[The Council Entity] will undergo [extensive turnover Event] [this year Time] [with some 21 of its 51 members stepping down Explanation].

Although frames and roles in FrameNet are very informative, because they are very fine-grained, their manual annotation is time-consuming and may convey more information than what is actually required within NewsReader. Therefore, we annotate participants information by adding a set of attributes to the *has_participant* relation that connects entity and event mentions. These attributes are inspired by Stanford typed dependencies with additional, generic information on the roles played by the fillers in the events. More details on this are described in Section 3.

2.3 Relations between entity mentions and between event mentions

2.3.1 Relations between entity mentions

Coreference relations link textual expressions, i.e. mentions, that refer to the same entity. One of the first annotation efforts on this topic dates back to the 6th and 7th Message Understanding Conferences (MUC) [Hirschman, 1997], where nouns, pronouns and noun phrases were considered as possible candidates, i.e. *markables*. During annotation, the headword of each markable expression was identified using the MIN attribute and its antecedent with the REF attribute as shown in the following example:

```
<COREF ID="1">Edna Fribble</COREF> and <COREF ID="2">Sam Morton</COREF>
   addressed the meeting yesterday.
<COREF ID="3" REF="1" TYPE="IDENT" MIN="Fribble">Ms. Fribble</COREF> discussed
   coreference, and
```

```
<COREF ID="4" REF="2" TYPE="IDENT" MIN="Morton">Mr. Morton</COREF> discussed
   unnamed entities.
```

The MUC approach to coreference has been criticized for mixing anaphora with other coreference phenomena [van Deemter and Kibble, 1995]. The same conflation is observed in the ACE (Automatic Content Extraction) program, in which annotators have to group mentions, whether named, nominal or pronominal, of the same entity together to perform intra-document coreference (see the latest version of the guidelines [Linguistic Data Consortium, 2008b]. In 2008, a cross-document global integration and reconciliation of information on annotation has been also performed but only for 50 person and organization entities and only for documents in which the target entities of interest were mentioned by name [Linguistic Data Consortium, 2008a]. Within the recent OntoNotes annotation, noun phrases, nominals (but not adjectival pre-modifiers) and verbs can be marked as coreferent [Technologies, 2011] but only in an intra-document perspective. In particular, two types of co-reference chains are marked, namely appositive constructions (e.g. *the PhacoFlex intraocular lens, the first foldable silicone lens available for cataract surgery*) and anaphoric coreference (e.g. *Elco Industries Inc. said it expects net income in the year ending June 30, 1990*).

In NewsReader, coreference relations between mentions within the same document are represented through REFERS_TO links between entity mentions and entity instances but also between event mentions and event instances. In this way, we partially follow the OntoNotes guidelines, because we consider also verbs as possible candidates (for details on event coreference see the following subsection while for more information on how to use the REFERS_TO link see Section 3.7). Given that coreferential relations are very language-specific, annotation will take into account the peculiarities of each language involved in the project: for example, for Spanish and Italian clitic pronouns will be properly handled. As far as cross-document coreference is concerned, NewsReader annotation is based on linking instances to external sources of information, such as DBpedia (`http://dbpedia.org`), following the two approaches described in Section 4.

2.3.2 Relations between event mentions

In NewsReader, we will consider three possible relations between events: coreference, temporal and causal relations.

Coreference relations: The MUC conferences [Grishman and Sundheim, 1996] in the '90s did not explicitly annotate events and **coreference** relations, but the templates used for evaluating the information extraction tasks indirectly can be seen as annotation of events represented in newswires. Such events are not ordered in time or further related to each other.

In ACE 2004 [Linguistic Data Consortium, 2004b], event detection and linking is included as a pilot task for the first time, inspired by annotation schemes developed for named entities. They distinguish between event mentions and the trigger event, which is the mention that most clearly expresses its occurrence [Linguistic Data Consortium, 2004a].

Typically, agreement on the trigger event is low across annotators (around 55% [Moens *et al.*, 2011]).

All these approaches have in common that they consider the textual representation as a closed world within which events need to be represented. This means that mentions are linked to a *trigger event* or to each other but not to an independent semantic representation. Also in TimeML event coreference is annotated as a temporal link between two event mentions, labeled as IDENTITY relation, even if strictly speaking coreference does not denote a temporal relation between two events. For this reason, in NewsReader we will model coreference as a different type of relation.

More recently, researchers started to annotate events across multiple documents, such as the *EventCorefBank* [Bejan and Harabagiu, 2010]. Cross-document coreference is more challenging also because it is not straightforward to establish which event in the chain should be the trigger event. Descriptions of events across documents may complement each other providing a more complete picture, but still textual descriptions tend to be incomplete and sparse with respect to time, place and participants. At the same time, the comparison of events becomes more complex. [Nothman *et al.*, 2012] define the trigger as the first new article that mentions an event, which is easier than to find the clearest description and still report inter-annotator agreement of .48 and .73, respectively. In the manual annotation of NewsReader documents, we tackle the problem of finding the trigger event in a completely different way: we lean on an external semantic representation of the event, which we call *event instance*, and we link each mention (intra- and cross-document) to it. This instance is possibly linked to DBPedia and is uniquely identified by time, place and participants. For each of such instances, a template is created in the *Cromer* tool for top-down event coreference.

Temporal relations: [Setzer and Gaizauskas, 2000] describe one of the first annotation frameworks to create coherent temporal orderings of events represented in documents using closure rules. They suggest that reasoning with text independent models, such as a calendar, helps annotating textual representations. In more recent years, TimeBank [Pustejovsky *et al.*, 2006b] has become the standard corpus for representing events and time-expressions. It has been annotated following the TimeML language [Pustejovsky *et al.*, 2003], in which temporal relations are marked via TLINKs. Each event (or time) is assigned a unique identifier, and these identifiers are used by TLINK annotations to assign one of the following temporal relations: BEFORE, AFTER, INCLUDES, IS_INCLUDED, DURING, DURING_INV, SIMULTANEOUS, IAFTER, IBEFORE, BEGINS, BEGUN_BY, ENDS or ENDED_BY. TimeML annotations were applied to the news stories in TimeBank, but agreement was low with annotators agreeing on which pairs of events and times to link only 55%, and agreeing on the relation type only 77% of the time. The following TempEval competitions [Verhagen *et al.*, 2007; Verhagen *et al.*, 2010] therefore tried to simplify the annotation scheme, annotating only temporal relations in certain syntactic constructions (e.g. the main events in adjacent sentences) and adopting a simpler relation set: BEFORE, AFTER, OVERLAP, BEFORE-OR-OVERLAP, OVERLAP-OR-AFTER and VAGUE. While requiring an explicit set of syntactic constructions resulted in 100% agreement on which pairs of events and times to tag, agreement on which temporal relation to assign was still low, between 65% and 72%. Anyway, during the last TempEval campaign ended in April 2013 [UzZaman *et al.*, 2013] the full set of TimeML temporal relations has been used instead of the coarse-grained version of previous editions. The same complete set will be adopted in NewsReader.

Causal relations: Annotation of causal relations is probably the most challenging task of WP3 because the connection between causes and effects is a complex phenomenon with roots in philosophy, psychology and linguistics and no consensus has been reached within the NLP community on a standard annotation scheme capturing causation.

Nevertheless, several attempts have been made to annotate causal relations in text. A common approach is to look for specific cue phrases such as "because" or "resulted" or to look for verbs that contain a cause as part of their meaning, such as "to break" (cause to be broken) or "to kill" (cause to die) [Khoo *et al.*, 2000; Sakaji *et al.*, 2008; Girju *et al.*, 2007]. Causal relations have also been annotated in the form of predicate-argument relations, for example in PropBank [Bonial *et al.*, 2010], where ARGM-CAU is used to annotate modifiers expressing the reason why the predicate takes place, for instance:

"The highway was [closed $_{Pred}$] [because of the snow$_{Argm-Cau}$]."

Finally, causal relations have been annotated as relations between events in a restricted set of linguistic constructions [Bethard *et al.*, 2008] and between clauses in text from novels [Grivaz, 2010]. The former define very simple annotation guidelines, asking annotators to add a causal relation between two events if the sentence(s) in which they appear can be paraphrased using a connective phrase such as "as a result" or "as a consequence". For instance, in the following examples, the first sentence should be annotated as containing a causal relation between events because it could be rephrased as shown in the second sentence:

"Fuel tanks had [leaked $_{Event1}$] and [contaminated $_{Event2}$] the soil."
⇒ "Fuel tanks had [leaked $_{Event1}$] and as a result [contaminated $_{Event2}$] the soil."

This approach is relatively simple for annotators, but agreement is only moderate (kappa of 0.556), in part because it focuses on events connected through "and", which is a highly ambiguous connective.

In [Grivaz, 2010], the paraphrasing approach presented by [Bethard *et al.*, 2008] is augmented with rules that combine linguistic tests with semantic reasoning tests. For instance, the annotation guidelines say that, given two events in a text, there is no causal relation between the two if the potential cause occurs after the potential effect. Besides, if the potential effect would probably have happened in the absence of the potential cause, the example is not causal. Despite the careful definition of the tests to perform, causal relations were still difficult to annotate, and high agreement (kappa of 0.84) was only achieved by the combined voting of four annotators.

Another approach to annotate causal relations has been introduced in the Penn Discourse TreeBank (PDTB) [The PDTB Research Group, 2008]. In this resource, relations are not annotated between specific event pairs but between two text spans called *Arguments*, which may correspond to sentences, clauses or event shorter segments. In PDTB, relations can be explicitly signaled by a set of lexically defined connectives (e.g. "because", "however"). In these cases, the relation is overtly marked, which makes it relatively easy to detect using NLP techniques. A relation between two discourse arguments, however, does not necessarily require an explicit connective, because it can be inferred also if a connective expression is missing. These cases, referred to as *implicit relations*, are annotated in PDTB only between adjacent sentences within paragraphs. If connective is not overt, PDTB annotators were asked to insert a connective to express the inferred relation.

Explicit: *The federal government suspended sales of U.S. savings bonds* <u>because</u> **Congress hasn't lifted the ceiling on government debt.**

Implicit: *The projects already under construction will increase Las Vegas's supply of hotel rooms by 11,795, or nearly 20%, to 75,500.* **By a rule of thumb of 1.5 new jobs for each new hotel room, Clark County will have nearly 18,000 new jobs**.

While in the first example above the connective "because" explicitly signals a causal relation holding between `Arg1` and `Arg2`, in the second no connective was originally expressed. A consequence relation is inferred between '*the increase in the number of rooms*' and '*the increase in the number of jobs*', though no *explicit* connective expresses this relation.

Each discourse relation is assigned a sense label based on a three-layered hierarchy of senses. The top-level, or *class level*, includes four major semantic classes, namely TEMPORAL, CONTINGENCY, COMPARISON and EXPANSION. For each class, a more fine-grained classification has been specified at *type* level. For instance, the relation in the explicit example above belongs to the CONTINGENCY class and the *Cause* type. A further level of *subtype* has been introduced to specify the semantic contribution of each argument.

At a more formal level, causal relations in PDTB belong to the CONTINGENCY:*cause* type and are identified when the situations described in `Arg1` and `Arg2` are causally influenced, but they are not in a conditional relation. Directionality is specified at the level of subtype with two different labels: "reason" (($\|\texttt{Arg2}\|<\|\texttt{Arg1}\|$[4]) and "result" ($\|\texttt{Arg1}\|<\|\texttt{Arg2}\|$) specifying which situation is the cause and which the effect. The typical connective for the first relation subtype is indeed *because*. On the contrary, for the latter (i.e. "result") , typical connectives are *so that*, *therefore*, *as a result*.

For cases in which there is no causal influence between `Arg1` and `Arg2`, because `Arg2` provides rather a justification for the claim expressed in `Arg1`, another type of cause has been introduced, called "Pragmatic Cause". We report an example (with implicit relation) below:

Implicit: *Mrs Yeargin is lying* [Implicit = because] **They found students in an advanced class a year earlier who said she gave them similar help**.

[4]The symbol $<$ used in the PDTB categories means "causes".

In PDTB, also temporal relations have been annotated, even though with a coarser-grained set of temporal relations compared to TimeML. The tag Temporal is used when the connective indicates that the situations described in the arguments are related temporally. Two types are defined for Temporal: *Asynchronous* (i.e., temporally ordered) and *Synchronous* (i.e., temporally overlapping). The first is used when the connective indicates that the situations described in the two arguments are temporally ordered. Two subtypes, i.e. "precedence" and "succession" are defined which specify whether it is `Arg1` or `Arg2` that describes an earlier event.

The tag *Synchronous* applies when the connective indicates that the situations described in `Arg1` and `Arg2` overlap. The type *Synchronous* does not specify the form of overlap like in TimeML, i.e., whether the two situations started and ended at the same time, whether one was temporally embedded in the other, or whether the two crossed. Typical connectives tagged as *Synchronous* are *while* and *when*.

2.4 Storylines

One of the goals of NewsReader is to capture the complexity of events in large amounts of news and to chain them according to temporal and causal relations. This task is particularly challenging because of the high complexity of news and the large amount of sources from which such data will be extracted.

Several theoretical works in the last forty years have proposed different models to capture event timelines and relations, mainly focusing on basic narrative structures like the ones found in stories. For instance, [Labov, 1972] defines *minimal narratives* as two states and a transition between them. In NewsReader, we may want to enrich it with the temporal, spatial and causal dimension, in order to distinguish between events taking place simultaneously in different locations and between events causing some other events. According to other theorists, narratives can be further analysed by distinguishing what is told, i.e. the actual content of a story, and the way it is told, i.e. how information is presented, what content is emphasized, etc. Although these two components have been named in different ways by different researchers, they are generally understood as the *fabula* and the *discourse* respectively (see [Gervás, 2010] for an overview focused on story generation). Recent attempts have been made in order to capture these information layers in a single computational framework for story understanding. [Lakoff and Narayanan, 2010], among others, have proposed a system for story analysis called KARMA that integrates the shared cognitive structures of human motivations, goals, emotions, actions, events and outcomes. The system is based on an ontological framework in which the rich structure of human event and action representation is encoded. Specifically, the ontology includes basic events defined by some parameters (e.g. the duration, the effects, the preconditions) and composite events composed of process primitives. The events are then connected through relations such as 'mutually exclusive', 'enable', 'subevent', etc. Although the authors report some problems that need to be tackled, for instance the identification of complex plot structures, their pilot study highlights the high number of dimensions involved in narratives and the effort required to fully analyse them.

Even though such approaches are very interesting, they are only partly suitable for NewsReader requirements. In fact, NewsReader foresees event extraction starting from large quantities of data, and a lightweight approach to event modelling is necessary to allow for an incremental enrichment of information on events, starting from the output of several NLP processors. However, the distinction between factual and non-factual events, as well as the point of view of the authors of the news and the other dimensions considered in KARMA will be taken into account both in NAF (i.e. the outcome of bottom-up annotation) and in SEM (i.e. the semantic model capturing events). As a starting point, we may consider the formal model represented in the News Storyline Ontology[5], a generic model for describing and organising storylines, which was specifically developed for capturing information coming from news streams.

The two annotation tasks described in Section 3 and 4 will provide the necessary building blocks for the more advanced building of storylines. To our knowledge, no tool exists for the manual annotation of storylines necessary to NewsReader. In this early stage of the project, no final decisions on how to model events and build storylines have been taken by the Consortium. Therefore, the annotation module will be built in the next project months and will be described in an updated version of this document (together with Deliverable D3.3.2).

3 Guidelines for annotation at document level

This section presents the annotation specifications defined in the project so far.

We propose to adopt a novel annotation format called NAF (NewsReader Annotation Format), which will be thoroughly described in D2.1. This format will be produced in output by the automatic NLP pipeline developed within the project but also by the tools for manual annotation. It is a standoff XML format in which several information layers can be easily added in an incremental way, and it is largely inspired by the Kyoto Annotation Format (KAF) [Bosma *et al.*, 2009] and the Terence Annotation Format (TAF) [Moens *et al.*, 2011].

For the annotation of events, time and participants at document level, our starting point is mainly TAF[6] because it was specifically defined to include event mentions, temporal expressions and participant mentions in a single annotation protocol. TAF is based on ISO-TimeML [Pustejovsky *et al.*, 2010], but introduces several adaptations in order to fit the domain of children's stories for which it was developed. The main novelty of TAF is related to the annotation of event participants. ISO-TimeML does not include the identification of event arguments, but the definition of the argument structure is essential to perform deep reasoning and full inference over events within texts. [Pustejovsky *et al.*, 2006a] address the need to include arguments in TimeML annotations, but that proposal does not include specific examples and details on how to perform annotation (e.g., on the

[5] `http://www.bbc.co.uk/ontologies/storyline/2013-05-01.html`

[6] See TERENCE European Project, ICT FP7 Programme, ICT-2010-25410 `http://www.terenceproject.eu/`

participants' attributes). Such guidelines were created for TAF and reused in NewsReader adapting them to the specific features of financial and economical texts.

TAF is compliant with NewsReader goals for four reasons. First, it incorporates the (in our opinion crucial) distinction between *instances* and *instance mentions*. Second, it adapts some consolidated paradigms for linguistic annotation such as TimeML for events and temporal expressions and ACE for participants and participant mentions [Linguistic Data Consortium, 2005]. It is thus compatible with other annotation schemes. Third, it integrates the annotation of event mentions, participants and temporal expressions into a unified framework. Fourth, it is complementary and compatible to KAF, to which it can be easily integrated.

As mentioned, TAF makes a clear distinction between *instances* and *instance mentions*. In the TERENCE project, this distinction only applied to nominal and named entities, similar to ACE [Linguistic Data Consortium, 2005], because children's stories, the main focus of that project, can generally be treated as a closed world, usually presenting a simple sequence of events that do not corefer. Event coreference and linking to other sources was thus not relevant for this domain. In NewsReader, we extend the distinction between instances and instance mentions to events in order to model event coreference, link them to other sources and create a consistent model for all instances. Details are given in the following subsections.

In the following subsections annotation specifications defined for the project are presented: they are independent from the formal annotation scheme (NAF) and from the semantic model (SEM) and strategic to train annotators. In general, instances will be represented in the semantic model while mentions will be represented in NAF.

3.1 Entities and Entity Mentions

Two different tags are used to distinguish between entity instances (i.e. <ENTITY>) and entity mentions (i.e. <ENTITY_MENTION>) in order to handle both the annotation of single mentions and of the coreference chains that link several mentions to the same entity in a text. Links between entity mentions and entity instances are annotated through a link named REFERS_TO (for more details see 3.7).

3.1.1 <ENTITY>

This tag is used to mark entities. An entity is an object or set of objects in the world or the mental representation of an object. It is classified from the semantic point of view, so the following semantic types have been defined:

1. PERSON, a single individual or a group of humans, e.g. *Barack Obama, a family*;
2. LOCATION, geographical regions both defined by political and/or social groups (e.g. *a nation*) or not (e.g. *landmasses and mountains*);
3. ORGANIZATION,corporations, agencies, and other groups of people defined by an established organizational structure, e.g. *Volkswagen*;

4. ARTIFACT, a human-made object, e.g. *a building, 500 cars*;
5. FINANCIAL, financial terminology widely used in NewsReader domain, e.g. *EGX-30 index, GDP*.

The original TAF specifications does not cover entities of type organization but, given their relevance in the financial domain, they have been added in the NewsReader annotation. Also the FINANCIAL entities have been added because they are essential to NewsReader domain. Each entity is described through an empty-element tag with the following attributes:

- id, automatically generated by the annotation tool;
- ent_type, it specifies the entity type from a semantic perspective. The possible values correspond to the 5 semantic types explained above;
- ent_class, it expresses the definiteness of the entity instance. The possible values are: SPC (Specific Referential), GEN (Generic Referential), USP (Under-specified Referential), NEG (Negatively Quantified).
- external_ref, it contains the URI used by an external source of information to identify a specific entity instance. This type of attribute would allow the representation of DBpedia entries and others;
- comment, a free text field where the annotator can add notes.

BNF of the ENTITY tag
attributes ::= id ent_type external_ref [comment]
id ::= <integer>
ent_type ::= PERSON | LOCATION | ORGANIZATION | ARTIFACT | FINANCIAL
ent_class ::= SPC | GEN | USP | NEG
external_ref ::= CDATA
comment ::= CDATA

3.1.2 <ENTITY_MENTION>

This tag encodes any textual realization of an entity, that is the portion of text in which an entity is referenced within a text. The extent of this portion of text is defined to be the entire nominal phrase used to refer to an entity, thus including modifiers (e.g. a big family), prepositional phrases (e.g. the President of the USA) and dependent clauses (e.g. John who is working in the garden). Annotators should tag all mentions of each entity within a document; for each mention, they identify the maximal extent of the string that describes the entity and label the head of the mention. Moreover, mentions are classified according to syntactic categories (e.g. proper names, common nouns, pronouns, etc.) given that entities may be referenced in a text by their proper name, indicated by a common noun or noun phrase, or represented by a pronoun.

Attributes of the <ENTITY_MENTION> tag are the following:

- id, automatically generated by the annotation tool;
- syntactic_type, it describes the syntactic category of the mention. The possible values are: NAM (proper name), NOM (nominal compound), PRO (pronoun), WHQ (question word), PTV (partitive), APP (appositional construction), CONJ (conjunction), PRE (pre-modifier), OTHER.
- comment.

BNF of the ENTITY_MENTION tag
attributes ::= id syntactic_type reference_type [comment]
id ::= <integer>
syntactic_type ::= NAM | NOM | PRO | WHQ | PTV | APP | CONJ | PRE | OTHER
comment ::= CDATA

In the sentence *'Qatar Navigation jumped 6.4 percent after the company said it scraped plans for a 20 percent capital increase'* the entity of type organization *Qatar Navigation* is expressed through 3 different textual realizations: *Qatar Navigation* is a mention of syntactic_type NAM, *the company* is a mention of syntactic_type NOM and *it* is a PRO mention.

3.2 Events and Event Mentions

In ISO-TimeML, *event* is used as a cover term to identify "something that can be said to obtain or hold true, to happen or to occur" [ISO TimeML Working Group, 2008]. In literature this notion is often referred to as eventuality [Bach, 1986] including all types of actions (punctuals or duratives) and states as well. Two different tags are adopted to distinguish between instances (i.e. <EVENT>) and instance mentions (i.e. <EVENT_MENTION>) of events in order to model event coreference.

3.2.1 <EVENT>

This tag is used to mark instances of events, that is the mental representations of events to which various types of linguistic elements (e.g. nouns, verbs, pronouns) refer within a text. Event instances are classified on the basis of 3 classes:

1. SPEECH_COGNITIVE, for speech acts and cognitive events, e.g. *report, say, think*;
2. CONTEXTUAL, events relevant for the financial domain, e.g. *buy, sell*;
3. GRAMMATICAL, e.g. *take place, hit, occur.*

Each event is described through an empty-element tag with the following attributes:

- id, automatically generated by the annotation tool;

- class, it specifies the event type, which values correspond to the 3 classes explained above;
- external_ref, it contains the URI used by an external source of information to identify a specific entity instance. This type of attribute would allow the representation of DBpedia entries and others;
- comment.

BNF of the EVENT tag
attributes ::= id class external_ref [comment]
id ::= <integer>
class ::= SPEECH_COGNITIVE | CONTEXTUAL | GRAMMATICAL | external_ref ::= CDATA comment ::= CDATA

3.2.2 <EVENT_MENTION>

This tag encodes different linguistic representations of a given event through a set of attributes largely inspired by those used in ISO-TimeML. Please note that the range of possible values of many attributes is great and depends on various grammatical and morphological features of particular languages: for details on attribute values see the English [Saurí *et al.*, 2010], Spanish [Saurí *et al.*, 2009] and Italian [Caselli, 2010] guidelines of TimeML annotation.

- id, automatically generated by the annotation tool;
- pred, it corresponds to the lemma of the token describing the event;
- factual, it conveys whether an event mention is presented as corresponding to a real situation in the world (YES) or to a situation that has not happened or has an uncertain status (NO);
- pos, it specifies the different grammatical categories which may realize an event, i.e. ADJECTIVE, NOUN, VERB, PREPOSITION, OTHER;
- tense, it captures standard distinctions in the grammatical category of verbal tense, i.e. PRESENT, PAST, FUTURE, NONE, INFINITIVE, PRESPART and PASTPART;
- aspect, it captures standard distinctions in the grammatical category of semantic aspect, i.e. NONE, PROGRESSIVE, PERFECTIVE, IMPERFECTIVE, PERFECTIVE_PROGRESSIVE, and IMPERFECTIVE_PROGRESSIVE;
- polarity, it distinguishes affirmative (POS) and negative (NEG) statements;

- mood, it identifies different modalities about the event realization, i.e. NONE, INDICATIVE, CONDITIONAL, SUBJUNCTIVE and IMPERATIVE. This attribute is used when syntactic mood is expressed by inflectional morphology on the verb thus it is to be annotated only in Italian and Spanish (in English and Dutch mood is expressed by means of modal auxiliary verbs);

- vform, it captures the distinctions between finite and non-finite verb forms, i.e. NONE, INFINITIVE, PARTICIPLE and GERUND. This attribute is optional in ISO-TimeML and it has been adopted in the Italian and Spanish guidelines but not in the English ones: because of this, this attribute will be annotated only for Italian and Spanish;

- modality, optional attribute that is used to convey different degrees of modality of an event, its value is the lemma of the modal verb modifying the main event, e.g. *may* (English), *potere* (Italian), *poder* (Spanish).

- comment.

The introduction of the *factual* attribute is a novelty with respect to TimeML in which the only characterization of event factuality is expressed through subordinating links (i.e. SLINKs) whereas in NewsReader we want to annotate the factuality degrees of event mentions [Roser and James, 2008] [Roser and James, 2009].

As for the identification of event mentions, the annotation of their extension is based on the notion of *minimal chunk*, because higher constituents may contain more than one event expression. This means that only the verbal head or the noun head is to be annotated in case of a VP or an NP construction, respectively, while event-denoting APs will have only their head adjective annotated as event.

BNF of the EVENT_MENTION tag

attributes ::= id pred pos factual tense aspect polarity [mood] [vForm] [modality] [comment]
id ::= <integer>
pred ::= CDATA
pos ::= ADJECTIVE | NOUN | VERB | PREPOSITION | OTHER
factual ::= YES | NO
tense ::= FUTURE | PAST | PRESENT | INFINITIVE | PRESPART | PASTPART | NONE
aspect ::= PROGRESSIVE | PERFECTIVE | IMPERFECTIVE | PERFECTIVE_PROGRESSIVE | IMPERFECTIVE_PROGRESSIVE |NONE
vform ::= INFINITIVE | GERUND | PARTICIPLE | NONE
polarity ::= NEG | POS
mood ::= INDICATIVE | CONDITIONAL | SUBJUNCTIVE | IMPERATIVE | NONE
modality ::= CDATA
comment ::= CDATA

3.3 Temporal Expressions

The <TIMEX3> tag taken from ISO-TimeML is used to annotate temporal expressions including both durations (e.g. three years) and points (e.g. *June 15th 2013, today*). Time points can be either absolute (e.g. *the 15th of June, 2013*) or underspecified expressions (e.g. *today*). Markable expressions can also be event anchored (e.g. *two days before the departure*) or sets of times (e.g. *every month*). The list of attributes selected for NewsReader and shown below is a reduced version of the list described in the ISO-TimeML guidelines:

- id, automatically generated by the annotation tool;
- value, it assigns a normalized value based on the ISO-8601 standard to the temporal expression. For example, the expression *June 15, 2013* would get the normalized form 2013-06-13 (YYYY-MM-DD), and the duration *60 days* would get the normalized form P60D (that means Period of 60 Days).
- type, it specifies the type of the temporal expression through 4 values, i.e. DATE, TIME, DURATION and SET.
- comment.

The original TAF does not adopt the *value* attribute because most children's stories are not fixed to a specific date therefore time normalization is often impossible: on the contrary, capturing the meaning of temporal expressions is crucial in news stories thus the attribute has been reintroduce in NewsReader.

BNF of the TIMEX3 tag
attributes ::= id type value [comment]
id ::= <integer>
value ::= CDATA
type ::= DATE | TIME | DURATION | SET
comment ::= CDATA

3.4 Numerical Expressions

In the MUC-7 Named Entity Task, monetary expressions and percentages were covered by a specific entity type named NUMEX[7]. The same type of numerical expressions were taken into account also in the Value Detection and Recognition task of the ACE program through the use of a dedicated markable called <VALUE> [?]. However, quantities in general are not explicitly covered neither by MUC annotation nor by ACE specifications. In ACE (and TAF) numerical quantities are annotated only if they are related to temporal expressions (e.g. 10 months) or if they are mentions of an entity (e.g. *VW had sold 400,000*

[7] http://cs.nyu.edu/faculty/grishman/NEtask20.book_17.html

cars: ***only 53,000*** *were gasoline-powered*) but not if they are entity pre-modifiers (as the quantity *400,000* in the previous sentence).

Given the relevance of numbers in the economic and financial domain, a new markable VALUE has been created for quantities (cardinal numbers in general) as well as for percentages and monetary expressions. The attribute *type* specifies if the expression is a percentage (e.g. *2.1 percent*), a capital described in terms of the currency of some country or region (e.g. *20 Euros*) or a quantity (e.g. *more than 500*). At the moment, this attribute is optional but it could become required in the second phase of the annotation.

BNF of the VALUE tag
attributes ::= id [type] [comment]
id ::= <integer>
type ::= PERCENT — MONEY — QUANTITY
comment ::= CDATA

3.5 Signals

The tag <SIGNAL>, inherited from ISO-TimeML, is used to annotate all those textual elements which make explicit a temporal relation (i.e. a TLINK, see 3.10) between two event mentions, two temporal expressions, or an event mention and a temporal expression. Annotators should identify temporal uses of prepositions (e.g. *during, from*), conjunctions (e.g. *before, after, while, when*) and adverbs (e.g. *meanwhile*) as well as mark up some special characters (e.g. the hyphen used in expressions denoting ranges, *2003-2005*). The tag contains only two attributes: id and comment.

BNF of the SIGNAL tag
attributes ::= id [comment]
id ::= <integer>
comment ::= CDATA

3.6 C-Signals

The <C-SIGNAL> tag has been introduced in TAF and consequently in NewsReader to mark-up textual elements that indicate the present of a causal relation (i.e. a CLINK, see 3.9). More specifically, annotators should identify all causal uses of conjunctions such as *because, thus, so* etc. The tag contains only two attributes: id and comment.

BNF of the C-SIGNAL tag
attributes ::= id [comment]
id ::= <integer>
comment ::= CDATA

3.7 Intra-document coreference

The relation REFERS_TO represents the coreference between an entity mention and an entity instance, and between an event mention and an event instance. It is a directional, many-to-one relation because many mentions can refer to the same entity (as in 1) or event (as in 2).

1. ***Qatar Navigation*** *jumped 6.4 percent after* ***the company*** *said* ***it*** *scraped plans for a 20 percent capital increase.*

2. *Indonesia's West Papua province was hit by a magnitude 6.1* ***earthquake*** *today, the latest powerful* ***tremor*** *to shake the region.*

The arguments taking part to the link are encoded into self-contained elements (i.e. <source .../>, <target .../>) and not annotated through specialized attributes as in standard TimeML, which uses `timeID` or `eventInstanceID` for the source of the relation and `relatedToTime` or `relatedToEventInstance` for the target. This solution allows for handling, if necessary many-to-one relations.

BNF of the REFERS_TO relation
attributes ::= id [comment]
id ::= <integer>
comment ::= CDATA

3.8 Participant Roles

The HAS_PARTICIPANT relation links an entity mention to the event mention it played a role in. More specifically, it is a directional, one-to-one relation from the event mention to the entity mentions describing the event participants. Note that for events other types of relations have been foreseen, i.e. TLINKs, CLINKs and SLINKs.

For asking questions about events reported in news, the most crucial relations are the subject and object relations. Besides, place and time are essential to uniquely identify an event. Thus, as a minimum requirement for NewsReader news, syntactic dependency relations between events and entities will be provided. This is acceptable as automatic dependency parsers exist for all project languages. Furthermore, FrameNet or PropBank-style semantic role parsers (and the extensive data required for training them) are only currently available in English. Thus in NewsReader, the annotation scheme will require dependency relations, while semantic role relations will be optional.

The choice of dependency labels to be assigned will be driven by the tagset of the parser(s) available for each language. As an indicative list, we suggest to adopt when possible the following minimal dependencies inspired by Stanford typed dependencies [de Marneffe and Manning, 2008a]:

- SUBJ: the syntactic subject of a clause.

- OBJ: the direct object of a VP.
- INDCOMPL: indirect complement of a VP.
- PREDCOMPL+SUBJ: predicative complement of the subject.
- PREDCOMPL+OBJ: predicative complement of the object.
- RMOD: modifier/adjunct.
- SUBJPass: subject of a passive sentence.
- INDCOMPLPass: prepositional complement with agent role in a passive construction.

We should take this list is only indicative, and not exhaustive. The actual number and type of relations will vary depending on requirements that we cannot totally foresee.

Also the semantic labels to be assigned to the annotated participants will strongly rely on the semantic paradigm adopted within NewsReader. A comparison between the performance of three available semantic role labellers (following FrameNet, PropBank and KYOTO framework) will be carried out in preparation of the first system version. In this context, the most suitable labels will be chosen also for manual annotation. In any case, the semantic type associated with the participants is already provided at entity level (e.g. person, location, organization, etc.).

BNF of the HAS_PARTICIPANT relation
attributes ::= id (ENG_dep | ES_dep | DU_dep | ITA_dep) sem-role
id ::= <integer>
ENG_dep ::= SUBJ | OBJ | INDCOMPL | PREDCOMPL+SUBJ | PREDCOMPL+OBJ | RMOD | SUBJPass | INDCOMPLPass | UNDEF
ES_dep::= SUBJ | OBJ | INDCOMPL | PREDCOMPL+SUBJ | PREDCOMPL+OBJ | RMOD | SUBJPass | INDCOMPLPass | UNDEF
DU_dep::= SUBJ | OBJ | INDCOMPL | PREDCOMPL+SUBJ | PREDCOMPL+OBJ | RMOD | SUBJPass | INDCOMPLPass | UNDEF
ITA_dep ::= SUBJ | OBJ | INDCOMPL | PREDCOMPL+SUBJ | PREDCOMPL+OBJ | RMOD | SUBJPass | INDCOMPLPass | UNDEF
sem_role ::= (FrameNet Set | PropBank Set | KYOTO Set)
comment ::= CDATA

3.9 Causal Relations Between Event Mentions

In NewsReader, we annotate causal relations between causes and effects denoted by event mentions through a link named CLINK. The format is analogous to the one of TLINKs. In a first phase, we do not distinguish between various type of causal relations, so CLINKS do not have a relType attribute. In case they will be distinguished at a later stage of the project, such distinctions will be introduced through a relType attribute.

When a causal relation is explicitly expressed by a conjunction, such as *by*, that conjunction is annotated with a C-SIGNAL tag (see Section 3.6) and its ID is reported in the c-signalID attribute of the CLINK relation. As for the other relations, also CLINKs have two self-contained elements to encode the source, that is the event mentions that were causes, and the target, that is the event mentions that were their effects, of the link.

BNF of the CLINK tag
attributes ::= id [c-signalID] comment
id ::= ID
c-signalID ::= IDREF
comment ::= CDATA

3.10 Temporal Relations

The TLINK relation conveys temporal links between pairs of event mentions, pairs of temporal expressions or between an event mention and a temporal expression. In order to create storylines, it is important to link each event with (at least) one other event in the text. There are 12 types of temporal relations, the same defined in ISO-TimeML except for the IDENTITY one given that in NewsReader coreferential relations are annotated by using the REFERS_TO link:

1. BEFORE, one event/time occurs before the other, e.g. *She **arrived** before his cousin **departure***;
2. AFTER, the opposite of BEFORE;
3. INCLUDES, one event/time includes the other, e.g. *John **left** on **Monday***;
4. IS_INCLUDED, the opposite of INCLUDES;
5. MEASURE, a duration expresses the persistence of an event, e.g. *John **ran** for **twenty minutes***;
6. SIMULTANEOUS, two events happen at the same time, e.g. *Mary was **watching** TV while John was **frying** the eggs*;
7. IBEFORE, one event/time occurs immediately before the other, e.g. *One of the eggs **crashed** as soon as it **touched** the pan*;
8. IAFTER, the opposite of IBEFORE;
9. BEGINS, a time or an event marks the beginning of another time or event, e.g. *From **morning** to twilight, John **drove** his car*;

10. BEGUN_BY, the opposite of BEGINS;
11. ENDS, a time or an event marks the ending of another event/time, e.g. *From morning to **twilight**, John **drove** his car*;
12. ENDED_BY, the opposite of ENDS.

For each relation, the following attributes are defined:

- id, automatically generated by the annotation tool;
- reltype, it specifies the type of the temporal relation through the 13 values described above;
- signalID, it represents the ID of the SIGNAL that explicitly signalled the presence of a TLINK.

As for the REFERS_TO link, also TLINKs have two self-contained elements to encode the source and the target of the link.

BNF of the TLINK tag
attributes ::= id [signalID] relType
id ::= ID
signalID ::= IDREF
signalID ::= SignalID
relType ::= BEFORE | AFTER | INCLUDES | IS_INCLUDED | SIMULTANEOUS | IAFTER | IBEFORE | BEGINS | ENDS | BEGUN_BY | ENDED_BY | MEASURE

3.11 Subordinating Relations

Annotation of reported speech or opinions is crucial in NewsReader because it is needed to understand the attitude of the writer / speaker with respect to what is being described. We annotate this leaning on TimeML approach, which uses SLINKs (subordination links) to connect REPORTING, I_STATE and I_ACTION verbs. In NewsReader, we reduce the scope of SLINKs by connecting through it event mentions belonging to the SPEECH_COGNITIVE type to their reported utterance/thought. In a second phase of the annotation, we will consider the possibility to mark the whole reported utterance, and not just the event mentioned in it, adopting an approach more similar to the one of PDTB [The PDTB Research Group, 2008] for attribution annotation. As for the other relations, also SLINKs have two self-contained elements to encode the source and the target of the link.

BNF of the SLINK relation
attributes ::= id [comment]
id ::= <integer>
comment ::= CDATA

3.12 Grammatical Relations

Sentences such as *the share drop comes on the same day* or *the earthquake hit at 7:48 am* contain an event-event relation between a nominal event (i.e. *drop* and *earthquake*) and another event, of type GRAMMATICAL, expressing its occurrence. To encode this type of relation, a GLINK relation has been defined in NewsReader annotation. As for the other relations, also GLINKs have two self-contained elements to encode the source and the target of the link.

BNF of the GLINK relation
attributes ::= id [comment]
id ::= <integer>
comment ::= CDATA

4 Guidelines for annotation at corpus level

The annotation at corpus level is introduced in NewsReader with the aim of marking-up coreference between entities and between events across different documents. While intra-document coreference is a well-established field of research at least for entities, the work on cross-document coreference is still burgeoning especially for events [Bejan and Harabagiu, 2010] [Lee *et al.*, 2012]. Our approach is to combine information taken from text with information taken from an external source of information (such as DBpedia) through a linking manually performed by annotators. The use of external sources of information allows us to correctly establish the fact that two or more expressions refer to the same entity or to the same event.

The annotation will be carried out following two approaches.

The first approach is **bottom-up**. Entity and event instances as well as mentions and relations between them will be marked in each text following the specifications described in Section 3. Cross-document coreference will be established by linking instances present in different texts taking the values given to the *DBpedia* attribute of entities and events as a reference.

The second approach is **top-down**. Here, a *seed set of entities and events* (i) of interest for the project, (ii) occurring at least in n_1 different documents of the NewsReader corpus, and (iii) occurring in no more than n_2 documents[8] will be selected creating a cluster of documents for each of them in an automatic way. The second constraint has been adopted to obtain interesting instances for cross-document coreference, whereas the third restriction will help us avoiding that annotation is too time consuming and error-prone. For each selected entity and event, a generic template to be filled in has been defined. Fields related to *entities* are the following:

[8] n_1 and n_2 are thresholds to be empirically set

- id, number that uniquely identifies the entity, automatically generated by the annotation tool;
- name, a human-friendly identifier of the entity;
- link, URI taken from an external knowledge base (e.g. DBpedia);
- type, corresponding to one of the 5 classes described in 3.1.

On the other hand, the following fields are assigned to each *event*:

- id, number that uniquely identifies the event, automatically generated by the annotation tool;
- name, a human-friendly identifier of the event;
- link, URI taken from an external knowledge base (e.g. DBpedia);
- date, it specifies the temporal dimension of the event;
- location, it specifies the spatial dimension of the event;
- participants, optional, lists entities that took part in the event;
- class, corresponding to one of the 3 classes described in 3.2.

Participants information in event templates are foreseen so as to link that field with the templates created for the corresponding entities at a later stage of the project.

While the kinds of entities we are going to annotate are likely to be present in DBpedia, events are less represented. Therefore, date, location and participants need to be specified if no DBpedia URI is given by the annotator, while they can be inherited from the URI if the annotator provides it.

In the upcoming months, some methodologies to speed up manual annotation will be tested and then implemented in the final annotation workflow. This will involve mainly *i)* the way document clusters are produced and *ii)* the way seed words are selected. For the first task, the purity of the clusters selected for entity and event annotation is crucial to reduce annotators' effort when annotating cross-document coreference. Several clustering methodologies will be tested based on document/topic similarity. Another important element to take into account is the need to include documents from different sources published at the same time and documents on the same topic published at different time points. The document creation time and the source information will be taken into account for this. As for the selection of seed words, also in this case it would be possible to test several methodologies to suggest a list of seed words that are relevant to a given cluster. This may be based, for instance, on keyword extraction. Furthermore, for each seed word (especially for events) an annotator may be provided with a list of possible synonyms, so as to make the search for the mentions in the documents as easy and complete as possible.

5 Using CAT for annotation at document level

The CELCT Annotation Tool (CAT) is a general-purpose web-based tool for text annotation developed by the Center for the Evaluation of Language and Communication Technologies (CELCT) [Bartalesi Lenzi *et al.*, 2012]. Its strong points are practical usability and customizability. In particular, CAT does not require programming skills nor prior knowledge about annotation tools and XML in order to install and use it. CAT has been successfully tested on TimeML annotation for the creation of part of the Ita-TimeBank, the largest Italian corpus annotated with information for temporal processing following the TimeML guidelines for Italian [Caselli *et al.*, 2011]. In addition, it is flexible enough to allow changes within an annotation schema: this feature is particularly useful at the beginning of a project when tagset definitions may change quite often. Because of these peculiarities, CAT has been identified as the tool that best meets the project requirements and that allows avoiding a time-consuming adaptation effort thus it will be used to perform annotation (for further details, see the following subsection 5.1). For the purposes of the project it will be possible to share the same account among different annotators allowing to monitor the work and perform collaborative annotation of the same files. In addition, a *NewsReader* annotation task will be created containing all the markables and relations defined for the project: this task will be shared among annotators and easily imported in all the accounts. Files annotated in CAT will be exported in a stand-off XML format: conversion module will be available to produce NAF from CAT native format and vice-versa. CAT is freely available for research purposes as a web application but we are going to release also a local CAT application creating an installation package. The package will provide a custom installation of Tomcat running on a non-standard port and, for data storage, this local application of CAT will use MySQL DataBase engine.

5.1 How to annotate with CAT

The tool interface (see Fig. 3) is based on four main components, namely the *Corpus Panel* (top left), where all corpora and files are listed, the *Layer Panel* (bottom left), where the annotation layers, the empty tags and the lexical categories are listed in different tabs, if present, the *Text Panel* (top right), with the text to be annotated, and the *Relation Panel* (bottom right), where relations are shown and can be created, modified or deleted, if necessary.

The objects to be annotated are called *Markables*, and correspond to a text string, either a single token or a sequence of words. In NewsReader, mentions of events and entities, temporal and numerical expressions, signals, c-signal and utterances are associated with a markable. For each of them, *attributes* can be defined using, for example, a radio button label, a text box or a drop-down menu. A screenshot of the graphical interface for the annotation of temporal expression attributes is displayed in Fig. 4.

CAT allows to annotate *empty tags*, that is tags with no textual content, for every type of markable defined for a task. In NewsReader, event and entity instances are encoded as empty tags and displayed in a dedicated tab of the Layer Panel, as shown in Fig. 5.

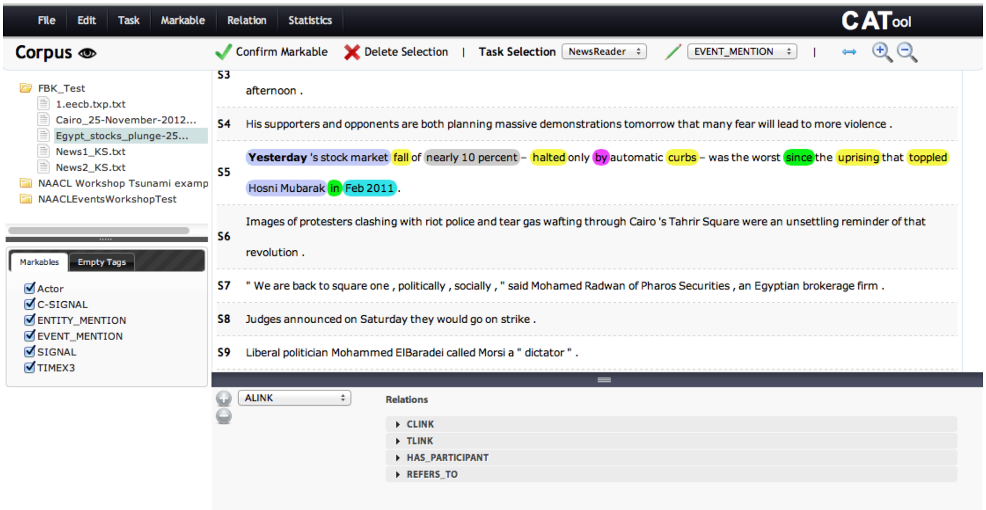

Figure 3: CAT interface

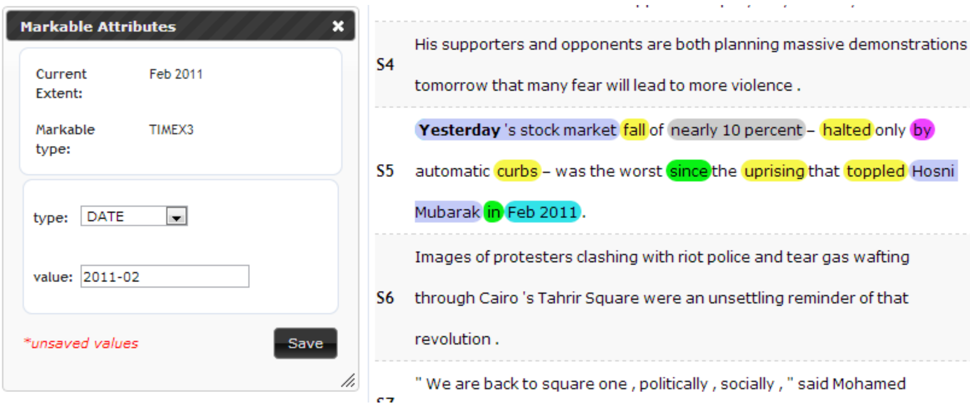

Figure 4: Example of CAT window for markable attributes

Figure 5: Example of CAT empty tags

Markables can be connected to each other by means of *relations*. In NewsReader there are 5 types of relations, namely TLINK, SLINK, GLINK, CLINK, HAS_PARTICIPANT, and REFERS_TO. A screenshot of the graphical interface for relation annotation is displayed in Fig. 6. In this case, a coreference is marked up through a REFERS_TO link between 3 mentions (i.e. *The kingdom's index, its, and its*) and the entity *Saudi Arabia*.

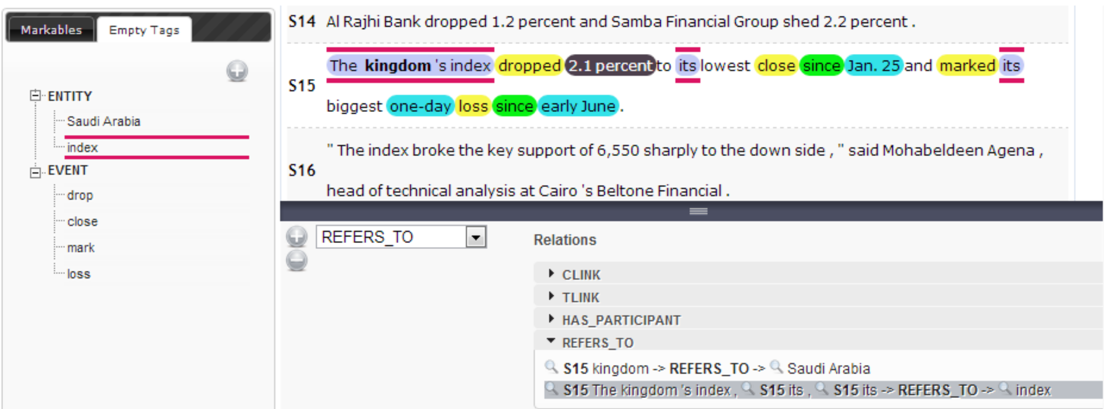

Figure 6: Annotation of relations

All the markables and relations defined for a project set up a task, in other words an annotation scheme, that can be exported and imported using an XML format.

6 Annotating cross-document coreference with CROMER

A multi-user web interface [Bentivogli *et al.*, 2008] specifically designed within the *Ontotext* project[9] for the cross-document coreference annotation task of named person entities will be extended to cope with the NewsReader requirements. The improved version of the tool is called CROMER (CRoss-document Main Event and entity Recognition). CROMER is composed of two pages, the Management Page and the Document Annotation Page. In the current version of the *Management Page* (see Fig. 7) annotators can look up a specific entity of type person in the corpus, retrieve the list of documents where it occurs and fill in the infobox template associated to it. The tool supports both exact string and partial string queries using the AND boolean operator or the wildcard "*". The box at the bottom on right side of the window contains all the entities created in correspondence with a given Seed Name and is used as entity repository during the document annotation process. If two or more different entities turn out to be the same, the "merge" button allows the annotator to merge them without having to annotate the documents again.

The new version of the interface will allow annotators to work on all the classes of entity instances defined in NewsReader (i.e. person, location, organization, artifact and financial) and on event instances as well to use the infobox templates described in Section 4. In the

[9] `http://ontotext.fbk.eu/`

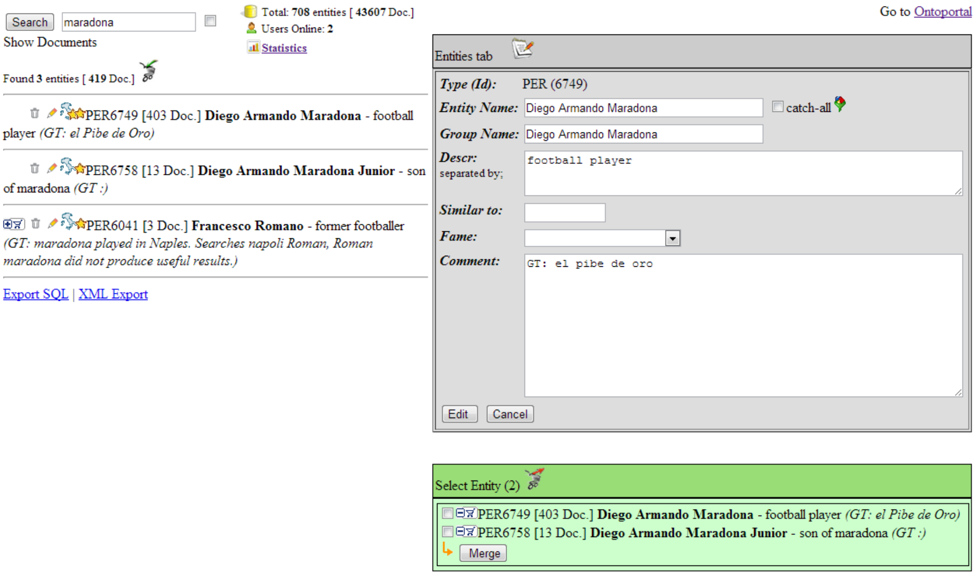

Figure 7: The Management Page of the current version of the tool.

Document Annotation Page (see Fig.8) the annotator submits a query and obtains all the documents satisfying the query, together with the text snippet in which the query string occurs. A scroll down menu is associated to each retrieved document, where the annotator can select the entity to which the document refers. The entities presented for annotation correspond to those inserted in templates created by the annotator in the Management Page. If the document snippets are not informative enough to individuate the correct entity, the annotator can also access the whole news.

Figure 8: The Document Annotation Page of the current version of the tool.

With the first version of the tool, annotators used to annotate the documents in which an entity is mentioned (in all its possible variants), but they could not identify the single mentions of the entity within the documents. This is due to the fact that the tool was first developed to perform cross-document annotation of named entities, based on the assumption that several occurrences of the same named entity in a document refer all to the same entity. In other words, if *Barack Obama* is mentioned several times in a document, it is enough to link such document to the template describing the current President of the United States without the need to explicitly link each mention in the document to such template.

While this assumption may still hold for entities annotated within NewsReader, it is unlikely to be valid for events. In fact, several mentions of an event can be present in a document, and it is very improbable that they should all refer to the same event instance. For example, if a 'purchase' event occurs several times in a financial news, it may describe different events that took place at a different place and time, with different participants. Therefore, when annotating events, CROMER will enable the linking of each event template to the cluster of documents where this event is described and then, in each document, *at least one mention* of such event will be marked. In this way, each event will also have a textual anchor. The other event mentions within the document referring to the same event can be recovered in a subsequent step either by matching this with the bottom-up annotation of the same document with CAT, or by running an event coreference system on the document. This feature allowing an explicit link between an event instance and a mention in a document has been introduced in order to minimize possible ambiguity issues. CROMER is freely available for research purposes as a web application: a stand-alone version of the tool will be released with a free license for research purposes as well, similar to CAT.

7 Conclusions

In this document, we describe the NewsReader annotation schemes for capturing information on events and the tools that will be used for this task. We give a detailed description of the annotation scheme and we provide a motivation of the choices made, trying to integrate existing standards with project-specific extensions. A first annotation test has been performed to take part to the shared task on events annotation, organized in the framework of the "First Workshop on Events"[10]. As described in the accompanying paper, annotation was largely based on the guidelines presented in this document, and the participating partners used a preliminary version of the CAT tool to annotate the shared task data without major issues.

Some minor adjustments of the annotation schemes are possible if they are necessary to meet the project requirements. A pilot annotation at document level on English data is already being performed, and the task description is reported in the deliverable appendix. A large scale annotation effort will start in month 9. Other extensions will concern the conversion modules, so that annotated data both with CAT and with CROMER will be exported in NAF format. However, we will need to have a final definition of NAF with all annotation layers in order to deliver such modules.

[10] `https://sites.google.com/site/cfpwsevents/`

8 APPENDIX - CAT Annotation Task for NewsReader

This appendix presents the first proposal of the CAT annotation task following the guidelines described in Section 3. The task is written in XML format and can be simply imported in the tool.

```
<?xml version="1.0" encoding="UTF-8"?>
<task name="NewsReader">
  <markables>
    <markable color="#f7f74a" name="EVENT_MENTION">
      <attributes>
        <attribute default_value="" name="pred" type="textbox" />
        <attribute default_value="YES" name="factual" type="radiobutton">
          <value value="YES" />
          <value value="NO" />
        </attribute>
        <attribute default_value="VERB" name="pos" type="combobox">
          <value value="VERB" />
          <value value="NOUN" />
          <value value="ADJECTIVE" />
          <value value="PREPOSITION" />
          <value value="OTHER" />
        </attribute>
        <attribute default_value="NONE" name="tense" type="combobox">
          <value value="PRESENT" />
          <value value="PAST" />
          <value value="FUTURE" />
          <value value="NONE" />
          <value value="INFINITIVE" />
          <value value="PRESPART" />
          <value value="PASTPART" />
        </attribute>
        <attribute default_value="PERFECTIVE" name="aspect" type="combobox">
          <value value="PROGRESSIVE" />
          <value value="PERFECTIVE" />
          <value value="IMPERFECTIVE" />
          <value value="NONE" />
          <value value="PERFECTIVE_PROGRESSIVE" />
          <value value="IMPERFECTIVE_PROGRESSIVE" />
        </attribute>
        <attribute default_value="POS" name="polarity" type="radiobutton">
          <value value="POS" />
          <value value="NEG" />
        </attribute>
        <attribute default_value="NONE" name="mood" type="combobox">
          <value value="NONE" />
```

```
        <value value="CONDITIONAL" />
        <value value="SUBJUNCTIVE" />
        <value value="IMPERATIVE" />
        <value value="INDICATIVE" />
      </attribute>
      <attribute default_value="NONE" name="vform" type="combobox">
        <value value="NONE" />
        <value value="INFINITIVE" />
        <value value="PARTICIPLE" />
        <value value="GERUND" />
      </attribute>
      <attribute default_value="" name="modality" type="textbox" />
      <attribute default_value="" name="comment" type="textbox" />
    </attributes>
  </markable>
  <markable color="#6b6b6b" name="ENTITY">
    <attributes>
      <attribute default_value="" name="ent_desc" type="textbox" />
      <attribute default_value="PERSON" name="ent_type" type="combobox">
        <value value="PERSON" />
        <value value="LOCATION" />
        <value value="ARTIFACT" />
        <value value="OTHER" />
        <value value="ORGANIZATION" />
        <value value="FINANCIAL" />
      </attribute>
       <attribute default_value="" name="head" type="textbox" />
      <attribute default_value="SPC" name="reference_type" type="combobox">
        <value value="SPC" />
        <value value="GEN" />
        <value value="USP" />
        <value value="NEG" />
      </attribute>
      <attribute default_value="" name="external_ref" type="textbox" />
    </attributes>
  </markable>
  <markable color="#c3caf7" name="ENTITY_MENTION">
    <attributes>
      <attribute default_value="NOM" name="syntactic_type" type="combobox">
        <value value="NAM" />
        <value value="NOM" />
        <value value="PRO" />
        <value value="WHQ" />
        <value value="PTV" />
        <value value="APP" />
```

```
        <value value="CONJ" />
        <value value="OTHER" />
        <value value="PRE" />
      </attribute>
      <attribute default_value="" name="comment" type="textbox" />
    </attributes>
  </markable>
  <markable color="#33e2e8" name="TIMEX3">
    <attributes>
      <attribute default_value="DATE" name="type" type="combobox">
        <value value="DATE" />
        <value value="TIME" />
        <value value="DURATION" />
        <value value="SET" />
      </attribute>
      <attribute default_value="" name="value" type="textbox" />
    </attributes>
  </markable>
  <markable color="#09f215" name="SIGNAL">
    <attributes>
      <attribute default_value="" name="comment" type="textbox" />
    </attributes>
  </markable>
  <markable color="#ef42ff" name="C-SIGNAL">
    <attributes>
      <attribute default_value="" name="comment" type="textbox" />
    </attributes>
  </markable>
  <markable color="#e31010" name="EVENT">
    <attributes>
      <attribute default_value="SPEECH_COGNITIVE" name="class" type="combobox"
         >
        <value value="SPEECH_COGNITIVE" />
        <value value="CONTEXTUAL" />
        <value value="GRAMMATICAL" />
      </attribute>
      <attribute default_value="" name="event_desc" type="textbox" />
      <attribute default_value="" name="external_ref" type="textbox" />
    </attributes>
  </markable>
  <markable color="#4d404d" name="VALUE">
    <attributes>
      <attribute default_value="" name="comment" type="textbox" />
      <attribute default_value="PERCENT" name="type" type="combobox">
        <value value="PERCENT" />
```

```
        <value value="MONEY" />
        <value value="QUANTITY" />
      </attribute>
    </attributes>
  </markable>
</markables>
<relations>
  <relation cardinality="one_to_one" color="#cbcbcb" direction="false" name="
      HAS_PARTICIPANT">
    <attributes>
      <attribute default_value="UNDEF" name="ENG_dep" type="combobox">
        <value value="UNDEF" />
        <value value="SUBJ" />
        <value value="OBJ" />
        <value value="INDCOMPL" />
        <value value="PREDCOMPL SUBJ" />
        <value value="PREDCOMPL OBJ" />
        <value value="RMOD" />
        <value value="SUBJPass" />
        <value value="INDCOMPLPass" />
      </attribute>
      <attribute default_value="UNDEF" name="ES_dep" type="combobox">
        <value value="UNDEF" />
        <value value="SUBJ" />
        <value value="OBJ" />
        <value value="INDCOMPL" />
        <value value="PREDCOMPL SUBJ" />
        <value value="PREDCOMPL OBJ" />
        <value value="RMOD" />
        <value value="SUBJPass" />
        <value value="INDCOMPLPass" />
      </attribute>
      <attribute default_value="UNDEF" name="DU_dep" type="combobox">
        <value value="UNDEF" />
        <value value="SUBJ" />
        <value value="OBJ" />
        <value value="INDCOMPL" />
        <value value="PREDCOMPL SUBJ" />
        <value value="PREDCOMPL OBJ" />
        <value value="RMOD" />
        <value value="SUBJPass" />
        <value value="INDCOMPLPass" />
      </attribute>
      <attribute default_value="UNDEF" name="ITA_dep" type="combobox">
        <value value="UNDEF" />
```

```
        <value value="SUBJ" />
        <value value="OBJ" />
        <value value="INDCOMPL" />
        <value value="RMOD" />
        <value value="PREDCOMPL SUBJ" />
        <value value="PREDCOMPL OBJ" />
        <value value="SUBJPass" />
        <value value="INDCOMPLPass" />
      </attribute>
      <attribute default_value="FrameNet Set" name="sem_role" type="combobox">
        <value value="FrameNet Set" />
        <value value="PropBank Set" />
        <value value="KYOTO Set" />
      </attribute>
      <attribute default_value="" name="comment" type="textbox" />
    </attributes>
  </relation>
  <relation cardinality="one_to_one" color="#cbcbcb" direction="false" name="
      TLINK">
    <attributes>
      <attribute default_value="" name="signalID" type="referenceLink" />
      <attribute default_value="" name="comment" type="textbox" />
      <attribute default_value="BEFORE" name="reltype" type="combobox">
        <value value="BEFORE" />
        <value value="AFTER" />
        <value value="IBEFORE" />
        <value value="IAFTER" />
        <value value="INCLUDES" />
        <value value="IS_INCLUDED" />
        <value value="MEASURE" />
        <value value="SIMULTANEOUS" />
        <value value="BEGINS" />
        <value value="BEGUN_BY" />
        <value value="ENDS" />
        <value value="ENDED_BY" />
        <value value="IDENTITY" />
      </attribute>
    </attributes>
  </relation>
  <relation cardinality="one_to_one" color="#cbcbcb" direction="false" name="
      CLINK">
    <attributes>
      <attribute default_value="" name="c-signalID" type="referenceLink" />
      <attribute default_value="" name="comment" type="textbox" />
    </attributes>
```

```
    </relation>
    <relation cardinality="many_to_one" color="#db1463" direction="true" name="
        REFERS_TO">
      <attributes>
        <attribute default_value="" name="comment" type="textbox" />
      </attributes>
    </relation>
    <relation cardinality="one_to_one" color="#cbcbcb" direction="false" name="
        SLINK">
      <attributes>
        <attribute default_value="" name="comment" type="textbox" />
      </attributes>
    </relation>
    <relation cardinality="one_to_one" color="#157528" direction="false" name="
        GLINK">
      <attributes>
        <attribute default_value="" name="comment" type="textbox" />
      </attributes>
    </relation>
  </relations>
</task>
```


References

[Bach, 1986] Emmon Bach. The algebra of events. *Linguistics and Philosophy*, 9:5–16, 1986.

[Bagga and Baldwin, 1998] Amit Bagga and Breck Baldwin. Algorithms for scoring coreference chains. In *In The First International Conference on Language Resources and Evaluation Workshop on Linguistics Coreference*, pages 563–566, 1998.

[Baker *et al.*, 2003] Collin F. Baker, Charles J. Fillmore, and Beau Cronin. The structure of the FrameNet database. *International Journal of Lexicography*, 16(3):281–296, 2003.

[Bartalesi Lenzi *et al.*, 2012] Valentina Bartalesi Lenzi, Giovanni Moretti, and Rachele Sprugnoli. CAT: the CELCT Annotation Tool. In *Proceedings of LREC 2012*, 2012.

[Bejan and Harabagiu, 2010] Cosmin Bejan and Sandra Harabagiu. Unsupervised event coreference resolution with rich linguistic features. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 1412–1422, 2010.

[Bentivogli *et al.*, 2008] Luisa Bentivogli, Christian Girardi, and Emanuele Pianta. In *Proceedings of the LREC 2008 Workshop on Resources and Evaluation for Identity Matching, Entity Resolution and Entity Management*, 2008.

[Bethard *et al.*, 2008] Steven Bethard, William Corvey, Sara Klingenstein, and James H. Martin. Building a corpus of temporal-causal structure. In European Language Resources Association (ELRA), editor, *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco, may 2008.

[Bonial *et al.*, 2010] Claire Bonial, Olga Babko-Malaya, Jinho D. Choi, Jena Hwang, and Martha Palmer. Propbank annotation guidelines, December 2010. `http://www.ldc.upenn.edu/Catalog/docs/LDC2011T03/propbank/english-propbank.pdf`.

[Bosma *et al.*, 2009] Wauter Bosma, Piek Vossen, Aitor Soroa, German Rigau, Maurizio Tesconi, Andrea Marchetti, Monica Monachini, and Carlo Aliprandi. KAF: a generic semantic annotation format. In *Proceedings of the 5th International Conference on Generative Approaches to the Lexicon GL 2009*, Pisa, Italy, 2009.

[Caselli *et al.*, 2011] Tommaso Caselli, Valentina Bartalesi Lenzi, Rachele Sprugnoli, Emanuele Pianta, and Irina Prodanof. Annotating Events, Temporal Expressions and Relations in Italian: the It-TimeML Experience for the Ita-TimeBank. In *Proceedings of LAW V*, Portland, Oregon, USA, 2011.

[Caselli, 2010] Tommaso Caselli. It-timeml: Timeml annotation scheme for italian, version 1.3.1, technical report. Technical report, Istituto di Linguistica Computazionale, Consiglio Nazionale delle Ricerche, Pisa, Italy, 2010. `http://www.ilc.cnr.it/pubblicazioni/ItTimeML_AnnotationGuidelines_1_3_1.pdf`.

[de Marneffe and Manning, 2008a] Marie-Catherine de Marneffe and Christopher D. Manning. Stanford typed dependencies manual, September 2008. `http://nlp.stanford.edu/software/dependencies_manual.pdf`.

[de Marneffe and Manning, 2008b] Marie-Catherine de Marneffe and Christopher D. Manning. The stanford typed dependencies representation. In *Coling 2008: Proceedings of the workshop on Cross-Framework and Cross-Domain Parser Evaluation*, pages 1–8. Association for Computational Linguistics, 2008.

[Ferro *et al.*, 2001] Lisa Ferro, Inderjeet Mani, Beth Sundheim, and George Wilson. Tides temporal annotation guidelines version 1.0.2, June 2001. `http://fofoca.mitre.org/annotation_guidelines/TIMEX2_2001_AnnotationGuide.pdf`.

[Ferro *et al.*, 2005] Lisa Ferro, Laurie Gerber, Inderjeet Mani, Beth Sundheim, and George Wilson. Tides 2005 standard for the annotation of temporal expressions, September 2005. `http://fofoca.mitre.org/annotation_guidelines/2005_timex2_standard_v1.1.pdf`.

[Gervás, 2010] Pablo Gervás. Computational approaches to storytelling and creativity. *AI Magazine*, 30(3):49–62, 2010.

[Girju *et al.*, 2007] Roxana Girju, Preslav Nakov, Vivi Nastase, Stan Szpakowicz, Peter Turney, and Deniz Yuret. Semeval-2007 task 04: Classification of semantic relations between nominals. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, pages 13–18, Prague, Czech Republic, June 2007. Association for Computational Linguistics.

[Grishman and Sundheim, 1996] Ralph Grishman and Beth Sundheim. Message understanding conference - 6: A brief history. In *Proceedings of the 16th conference on Computational linguistics (COLING'96)*, pages 466–471, 1996.

[Grivaz, 2010] Cécile Grivaz. Human judgements on causation in french texts. In Nicoletta Calzolari (Conference Chair), Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, Mike Rosner, and Daniel Tapias, editors, *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10)*, Valletta, Malta, may 2010. European Language Resources Association (ELRA).

[He, 2007] Tian He. *Coreference Resolution on Entities and Events for Hospital Discharge Summaries*. PhD thesis, Massachusetts Institute of Technology, 2007.

[Hirschman, 1997] Lynette Hirschman. MUC-7 coreference task definition. In *Proceedings of the 7th Message Understanding Conference*, 1997. `http://www.itl.nist.gov/iaui/894.02/related_projects/muc/proceedings/co_task.html`.

[ISO TimeML Working Group, 2008] ISO TimeML Working Group. ISO TC37 draft international standard DIS 24617-1, August 14 2008. `http://semantic-annotation.uvt.nl/ISO-TimeML-08-13-2008-vankiyong.pdf`.

[Khoo *et al.*, 2000] Christopher S. G. Khoo, Syin Chan, and Yun Niu. Extracting causal knowledge from a medical database using graphical patterns. In *In Proceedings of 38th Annual Meeting of the ACL, Hong Kong, 2000*, pages 336–343, 2000.

[Labov, 1972] W. Labov. *Language in the Inner City: Studies in the Black English Vernacular.* University of Pennsylvania Press, 1972.

[Lakoff and Narayanan, 2010] G. Lakoff and S. Narayanan. Toward a computational model of narrative. In *AAAI Fall Symposium Series*, 2010.

[Lee *et al.*, 2012] Heeyoung Lee, Marta Recasens, Angel Chang, Mihai Surdeanu, and Dan Jurafsky. Joint entity and event coreference resolution across documents. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pages 489–500, Jeju Island, Korea, July 2012. Association for Computational Linguistics.

[Linguistic Data Consortium, 2004a] Linguistic Data Consortium. Annotation Guidelines for Event Detection and Characterization (EDC). `http://projects.ldc.upenn.edu/ace/docs/EnglishEDCV2.0.pdf`, 2004.

[Linguistic Data Consortium, 2004b] Linguistic Data Consortium. The ACE 2004 Evaluation Plan. Technical report, LDC, 2004.

[Linguistic Data Consortium, 2005] Linguistic Data Consortium. ACE (Automatic Content Extraction) English annotation guidelines for entities. Version 6.6, July 2005.

[Linguistic Data Consortium, 2008a] Linguistic Data Consortium. Ace 2008: Cross-document annotation guidelines (xdoc), version 1.6 – 2008.05.06. Technical report, 2008. `http://projects.ldc.upenn.edu/ace/docs/ACE08_XDOC_1.6.pdf`.

[Linguistic Data Consortium, 2008b] Linguistic Data Consortium. Ace (automatic content extraction) english annotation guidelines for entities, version 6.6 2008.06.13. Technical report, June 2008. `http://projects.ldc.upenn.edu/ace/docs/English-Entities-Guidelines_v6.6.pdf`.

[Marcus *et al.*, 1994] Mitchell Marcus, Grace Kim, Mary Ann Marcinkiewicz, Robert MacIntyre, Ann Bies, Mark Ferguson, Karen Katz, and Britta Schasberger. The penn treebank: annotating predicate argument structure. In *Proceedings of the workshop on Human Language Technology*, HLT '94, pages 114–119, Stroudsburg, PA, USA, 1994. Association for Computational Linguistics.

[Moens *et al.*, 2011] Marie-Francine Moens, Oleksandr Kolomiyets, Emanuele Pianta, Sara Tonelli, and Steven Bethard. D3.1: State-of-the-art and design of novel annotation languages and technologies: Updated version. Technical report, TERENCE project – ICT FP7 Programme – ICT-2010-25410, 2011.

[Nothman *et al.*, 2012] Joel Nothman, Matthew Honnibal, Ben Hachey, and James R. Curran. Event linking: Grounding event reference in a news archive. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 228–232, Jeju Island, Korea, July 2012. Association for Computational Linguistics.

[Palmer *et al.*, 2005] Martha Palmer, Daniel Gildea, and Paul Kingsbury. The proposition bank: An annotated corpus of semantic roles. *Computational Linguistics*, 31(1):71–106, 2013/03/12 2005.

[Pustejovsky *et al.*, 2003] J. Pustejovsky, J. Castano, R. Ingria, R. Saurí, R. Gaizauskas, A. Setzer, G. Katz, and D. Radev. Timeml: Robust specification of event and temporal expressions in text. In *IWCS-5 Fifth International Workshop on Computational Semantics*, 2003.

[Pustejovsky *et al.*, 2006a] James Pustejovsky, Jessica Littman, and Roser Saurì. Argument Structure in TimeML. In *Dagstuhl Seminar Proceedings. Internationales Begegnungs- und Forschungszentrum*, 2006.

[Pustejovsky *et al.*, 2006b] James Pustejovsky, Jessica Littman, Roser Saurí, and Marc Verhagen. Timebank 1.2 documentation. Technical report, Brandeis University, April 2006.

[Pustejovsky *et al.*, 2007] J. Pustejovsky, J. Littman, and R. Saurí. Arguments in timeml: events and entities. *Annotating, Extracting and Reasoning about Time and Events*, pages 107–126, 2007.

[Pustejovsky *et al.*, 2010] James Pustejovsky, Kiyong Lee, Harry Bunt, and Laurent Romary. ISO-TimeML: An international standard for semantic annotation. In *Proceedings o the Fifth International Workshop on Interoperable Semantic Annotation*, 2010.

[Roser and James, 2008] Saurí Roser and Pustejovsky James. From structure to interpretation: A double-layered annotation for event factuality. In *Prooceedings of the 2nd Linguistic Annotation Workshop*, 2008.

[Roser and James, 2009] Saurí Roser and Pustejovsky James. Factbank: A corpus annotated with event factuality. *Language Resources and Evaluation*, 2009.

[Sakaji *et al.*, 2008] Hiroki Sakaji, Satoshi Sekine, and Shigeru Masuyama. Extracting causal knowledge using clue phrases and syntactic patterns. In *Proceedings of the 7th International Conference on Practical Aspects of Knowledge Management*, PAKM '08, pages 111–122, Berlin, Heidelberg, 2008. Springer-Verlag.

[Saurí *et al.*, 2009] Roser Saurí, Olga Batiukova, and James Pustejovsky. Annotating events in spanish, technical report bm 2009-01. 2009. `http://www.ldc.upenn.edu/Catalog/docs/LDC2012T12/annotationGuidelines_event_SP.pdf`.

[Saurí *et al.*, 2010] Roser Saurí, Lotus Goldberg, Marc Verhagen, and James Pustejovsky. Annotating events in english, technical report 2010-05. 2010. `http://www.timeml.org/tempeval2/tempeval2-trial/guidelines/EventGuidelines-050409.pdf`.

[Setzer and Gaizauskas, 2000] Andrea Setzer and Robert J. Gaizauskas. Annotating events and temporal information in newswire texts. In *LREC*. European Language Resources Association, 2000.

[Technologies, 2011] BBN Technologies. Co-reference guidelines for english ontonotes version 6.0, 2011. `http://www.ldc.upenn.edu/Catalog/docs/LDC2011T03/coreference/english-coref.pdf`.

[The PDTB Research Group, 2008] The PDTB Research Group. The PDTB 2.0. Annotation Manual. Technical Report IRCS-08-01, Institute for Research in Cognitive Science, University of Pennsylvania, 2008.

[UzZaman *et al.*, 2013] Naushad UzZaman, Hector Llorens, Leon Derczynski, Marc Verhagen, James Allen, and James Pustejovsky. Semeval-2013 task 1: Tempeval-3: Evaluating events, time expressions, and temporal relations. In *Proceedings of the 7th International Workshop on Semantic Evaluation (SemEval 2013)*, 2013.

[van Deemter and Kibble, 1995] Kees van Deemter and Rodger Kibble. On coreferring: Coreference in muc and related annotation schemes. *Computational Linguistics*, 26:629–637, 1995.

[Verhagen *et al.*, 2007] Marc Verhagen, Robert Gaizauskas, Frank Schilder, Mark Hepple, Graham Katz, and James Pustejovsky. Semeval-2007 task 15: Tempeval temporal relation identification. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, pages 75–80, Prague, Czech Republic, June 2007. Association for Computational Linguistics.

[Verhagen *et al.*, 2010] Marc Verhagen, Roser Saurrí, Tommaso Caselli, and James Pustejovsky. Semeval-2010 task 13: Tempeval-2. In *Proceedings of the 5th International Workshop on Semantic Evaluation*, pages 57–62, Uppsala, Sweden, July 2010. Association for Computational Linguistics.

[Weischedel *et al.*, 2010] Ralph Weischedel, Sameer Pradhan, Lance Ramshaw, Jeff Kaufman, Michelle Franchini, Mohammed El-Bachouti, Nianwen Xue, Martha Palmer, Mitchell Marcus, Ann Taylor, Craig Greenberg, Eduard Hovy, Robert Belvin, and Ann Houston. Ontonotes release 4.0, December 2010. `http://www.ldc.upenn.edu/Catalog/docs/LDC2011T03/OntoNotes-Release-4.0.pdf`.