

Events in the Vaccination Corpus Annotation Guidelines

VUA Perspectives Group

January 23, 2018

Contact person: Roser Morante (r.morantevallejo@vu.nl)

Contents

1	Event: A Definition	3
2	Event Annotation in General	3
2.1	What is an event?	4
2.2	When are states to be annotated as events?	4
2.3	Implicit event readings	4
3	<EVENT>: tag extent	5
4	Events realized by VERBS	5
5	Events realized by NOUNS	6
6	Events realized by ADJECTIVES	8
7	Events realized by PREPOSITIONAL PHRASES	8
8	Diverse constructions	9
8.1	Light verb constructions	9
8.2	Copulative constructions	9
8.3	Causative constructions	10
9	Event Annotation in the Vaccination Corpus	10
9.1	Arguments in favour and against	10
9.2	Events that should be annotated	11
9.3	Events that should not be annotated	12
10	Annotation of specific cases	13
10.1	Vaccine, vaccination	13
10.2	Diseases	14
10.3	Other	14
11	Example text	14

12 Annotation in eHOST: Instructions for Annotators	15
12.1 Configuring the Workspace	15
12.2 Introduce annotator name	17
12.3 Selecting a project	18
12.4 Annotating markables	19
12.5 Visualising already annotated markables	19
12.6 Discontinuous span for markables	20
12.7 Adjusting the span of a markable	21
12.8 Saving your data	22

This document contains the guidelines that will be used to annotate the Vaccination Corpus with events that are related to the vaccination debate. Before starting the annotation you should be familiar with several concepts:

- What is an event.
- How should events be annotated in general.
- What is the vaccination debate about.
- How to determine whether an event is related to the vaccination debate.

These guidelines are based on the TimeML Annotation Guidelines (Pustejovsky et al., 2006), the News-Reader Annotation Guidelines (Tonelli et al., 2014), the ECB+ Annotation Guidelines (Cybulska and Vossen, 2014), the ISO-TimeML Framework (ISO, 2008), and the Richer Event Description (RED) Guidelines v1.7¹ (Styler et al., 2014). The guidelines apply to English and at document and mention levels.

These guidelines provide the information that you need. Section 1 defines what is an event. Section 2 provides annotation guidelines for events in general. Section 9 explains which events are related to the vaccination debate. It is important to remember that this task is not about annotating all events in the corpus, but only events that are related to the vaccination debate. Finally, Section 12 presents details about how to annotate with the annotation tool eHost.

1 Event: A Definition

TimeML (and ISO-TimeML) defines events as “situations that happen or occur” (Pustejovsky et al., 2006). In TimeML, *event* is used as a cover term for describing both dynamic, static, durative and non-durative situations. States are additionally defined as “circumstances in which something obtains or holds true” (ibid.).

In RED an event is defined as “any occurrence, action, process or event state which deserves a place upon a timeline, and could have any syntactic realization”. This latter definition, based on semantic considerations and independent of syntactic considerations, will be adopted in the remainder of the guidelines as our definition of what is an event.

However, you should bear in mind, that, as described in Section 9, for this task we only annotate events that are related to the vaccination debate.

2 Event Annotation in General

The tag `<EVENT>` is used to annotate all linguistic realizations of (temporally relevant) event mentions.²

¹<https://github.com/timjogorman/RicherEventDescription/blob/master/guidelines.md>

²Note that in the (ISO, 2008) the term “instance” is used to refer to event mentions.

2.1 What is an event?

Following the RED guidelines, the first task is to distinguish between real event mentions and mentions which encode relationships (e.g. causal or aspectual, among others) or correspond to purely grammatical elements (e.g. auxiliaries, modals, light verbs, and support verbs, among others). The key questions to be asked are:

- Is a certain group of lexical items or a single one expressing something that is happening, happened, will or may happen?
- Is a certain group of lexical items or a single one expressing a change of state, a process or of a set of processes, a situation, or a temporally bounded property?

In the following (and the rest of) examples, events are underlined:

1. *People are buzzing about the release of the new iPhone.*
2. *I might come with you if you don't mind.*
3. *You seem to be happy.*
4. *The earthquake hit the coast*
5. *The party started at 10 o'clock.*

2.2 When are states to be annotated as events?

As reported in the RED guidelines, “all states **exist** on a timeline”, thus this criterion is only necessary but not sufficient to decide which states constitute an event or not. The annotation of states is restricted to **temporally bounded states**, thus **excluding permanent properties** (e.g. being tall or short in a person). In addition to this, we will adopt a further criterion: a state is considered as an event when “its use implies actual occurrences such as the events leading up to its own existence.”³. Consequently, in Example (6) we annotate an event, whereas in Example (7) we don't.

6. *I came home and the door was open.* → it implies an event of opening the door.
7. *He walked through the open door.* → it does not imply an event of opening the door.

Furthermore, adjectives used as specifiers or mere modifiers of entities (i.e. to restrict the description of an object or person) should not be annotated as event (e.g. “the *tall* boy”, “the *yellow* canary”, “the *retracted* study”).

2.3 Implicit event readings

Some occurrences of event mentions may be complicated to identify as they may correspond/denote or seem to denote more than one entity at the same time:⁴

³<https://github.com/timjogorman/RicherEventDescription/blob/master/guidelines.md#what-is-an-event>

⁴The following examples are taken from the RED Guidelines v1.7

8. *Remember to eat before the vitamins.*
9. *I recommend Tylenol three times a day.*
10. *He was sentenced to five years of prison.*
11. *A bomb disrupted the parade.*

These examples illustrate cases where the linguistic elements that are underlined assume an implicit event reading: something that happens in a location (the state of being implied by “*prison*”), a typical action (consuming/taking pills implied by “*vitamins*”; explosion implied by “*bomb*”), or a coreference relation with a “prototypical” event mention. Cases of implicit events must be annotated always.

3 <EVENT>: tag extent

Following the surface-oriented TimeML annotation philosophy, event annotation is based on the notion of *minimal chunks*, because higher constituents may contain more than one event expression. This means that **only the head of the event denoting chunk will be marked up with the tag. Auxiliaries, polarity markers, particles, modifiers, complements and specifiers are not part of the extent of the tag.** Nevertheless, to be more informative at the semantic layer, the minimal chunk rule is applied in a more flexible manner.

Multitoken event mentions are allowed though restricted to:

- Idioms. the whole idiom is to be annotated if it appears as an entry in the British or American version of the Collins English Dictionary online.⁵

12. *My computer has kicked the bucket. (idiom)*

- Proper nouns
- Prepositional phrases

Multitoken events are always annotated with a unique <EVENT> tag, even when their surface realization is discontinuous.

4 Events realized by VERBS

All verbal predicates, including those that denote a state, give rise to events. Both finite and non-finite verb forms are annotated. The event token corresponds to the verbal head of the VP. Auxiliaries (both temporal and modal) are excluded from the tag extent;

13. *People are buzzing about the release of the new iPhone.*
14. *I might come with you if you don't mind.*

⁵<http://www.collinsdictionary.com/dictionary/english/>

Mentions of **aspectual verbs** (e.g. start, begin, continue, stop, finish, initiate, re-initiate, complete, etc.) are not to be annotated when they contribute to signaling grammatical information concerning the BOUNDEDNESS of an event; e.g.:

15. *We completed treatment today.* → “completed” is an aspectual event, no annotation.
16. *We completed the form.* → “completed” is a proper event
17. *Measles has made a comeback.* → “to make a comeback” is like re-initiate.

As for phrasal verbs, only the verbal head is annotated:

18. *He switched the light off.*

5 Events realized by NOUNS

Events realized by nouns can correspond to:

- nominalizations;
- event nouns;
- contextual event readings or implicit events.

With “contextual event reading” of nouns we refer to a set of well known semantic phenomena, such as type-coercion (e.g. *beginning a book*), metonymy (e.g. *The bomb ended the manifestation.*) and similar, which elicit an eventive reading of a noun which would normally refer to an entity (either concrete or abstract) or imply the occurrence of an event.

19. *The flight was scheduled at 08.00 p.m.*
20. *The meeting lasted 3 hours.*
21. *A panic attack can be dangerous.*
22. *An allergic reaction can be dangerous.*
23. *A bomb disrupted the parade.*

An extension with respect to TimeML and ISO-TimeML is the annotation of events realized by proper nouns (e.g. *World War II*). For this kind of events, the entire noun phrase realizing the event description must be annotated with a single <EVENT> tag, including temporal or other kinds of modifiers both in pre-nominal and post-nominal positions.

24. *The 1972 Summer Olympics ended in tragedy.*
- 25.
26. *Then he got Hashimoto’s disease.*

Event-denoting nouns in pre-modifier position are never to be annotated as event mentions. This is why *election*, *panic* and *recision biopsy* are not annotated in the examples below.

27. *The election defeat was annoying.*
28. *The panic attack was scaring.*
29. *That recision biopsy analysis of the sigmoid colon today.*

The same applies for present participles in pre-modifier position:

30. *recurring event*
31. *drinking problem*

As for nouns occurring in constructions of the kind “ N_1 of N_2 ”, the following cases can occur:

- Both nouns refer to entities: no element must be annotated;
32. *The cat_{N1} of my mother_{N2}.*
- N_1 functions as an aspectualizer of N_2 (i.e. it makes explicit grammatical information related to BOUNDEDNESS, ABRUPTNESS and DURATION (Simone and Masini, 2007)), then N_1 must NOT be marked with the <EVENT> tag; only N_2 must be annotated if an event;
33. *the begin_{N1} of the plague_{N2}.*
 34. *a burst_{N1} of enthusiasm_{N2}.*
 35. *the outbreak_{N1} of measles_{N2}.*
- N_1 denotes an event but N_2 does not, then only N_1 will be marked with the <EVENT> tag;
36. *the shot_{N1} of the gun_{N2}.*

Similarly to verbs, aspectual nouns are not annotated (as already stated above) as events. Nevertheless, some nouns which may function as an aspectualizer of an event (e.g. “outbreak”) must be annotated if subsequently in the text they are (coreferentially) referred to as proper events. Consider this example:

37. *The outbreak of the measles was unexpected.*
38. *The outbreak of the measles was unexpected. The outbreak occurred yesterday in a school downtown.*
39. *The emergence of these diseases.*

In some cases the noun that provides aspectual information provides also additional lexical information. In these case we annotate the noun as event.

40. *The elimination of measles.*

6 Events realized by ADJECTIVES

Adjectives normally denote stative events. Adjectives are to be annotated only when they denote temporally bounded properties or when their presence implies the actual occurrence of another event. When in predicative position, only the adjectival head of the predicative construction must be annotated. What follows are examples of adjectives as predicative complements of verbal constructions:

- Copular constructions (*be*, *seem*, etc.);

41. *The crowd seems angry.*

- Inchoative predicates;

42. *The crowd became angry.*

43. *The vaccine became available.*

- Aspectual predicates;

44. *The crowd keeps angry.*

- Causative predicates;

45. *The police made the crowd angry.*

- Perception predicates;

46. *The crowd looked angry.*

- Evaluative and descriptive predicates.

47. *The crowd is often described as angry.*

The list above is not exhaustive. Please remember that:

- All adjectives in predicative position must be annotated as events even if they are NOT the complement of a verbal predicate but temporally bounded;
- In case of doubt DO NOT mark the adjective.

As for adjectives in attributive position, they can be marked as events only when their occurrence “implies actual occurrences such as the events leading up to its own existence”. In addition to this, in case of a clear coreference relation or SET/MEMBER relationship with other adjectives marked as event, adjectives must be marked as event if they fail the implied event test.

7 Events realized by PREPOSITIONAL PHRASES

Prepositional Phrases (PPs) denote an event only when in predicative constructions, i.e. predicative complements (e.g. copular constructions, inchoative predicates, aspectual predicates, and causative predicates). Only the head of the prepositional phrase (i.e. the preposition) is to be annotated.⁶ However a few exceptions to this rule apply:

⁶Good examples to be found.

- in case the PP is an entry in the British or American version of the Collins English Dictionary online,⁷ the minimal chunk rule does not apply and the extent of the tag corresponds to the whole expression.
48. *According to the president you must resign.* (*According to* implies that someone has said/believes something.)
49. *All people on board died.* (*On board* implies that someone is on board.)
- If the verb, noun, or adjective head of the PP denotes an event itself, then the verb, noun, or adjective head must be annotated;
50. *The program started with an interview.*
51. *He is characterized as eccentric.*

8 Diverse constructions

8.1 Light verb constructions

Light verb constructions are predicates involving a verb with null or very low semantic content (e.g. *make*, *get*, *take*, *put*, *have*, etc.) and a noun as its argument. Only the noun will be annotated.

52. *He took a shower.*
53. *I got more support from my boss.*
54. *Most people do get their vaccinations.*
- 55.
56. *Then he got Hashimoto's disease.*

8.2 Copulative constructions

Copulative constructions consist of a predicate realized by a copular verb (*be*, *seem* etc.) followed by an NP, AP or PP. Only the predicative complement will be annotated following the rules for NP, PP or AP annotation:

57. *Hillary Clinton is the most successful candidate for the 2016 election.* (*the most successful candidate for the 2016 election* is an entity, not an event.⁸)
58. *There will be deaths.*
59. *The crowd looked angry.* (*looked* is a perception predicate.)
60. *A disease that is highly contagious may kill people.*

⁷<http://www.collinsdictionary.com/dictionary/english/>

⁸Sortal states are not annotated.

8.3 Causative constructions

Causative constructions are normally composed by a triplet: the logical subject, the causal expression and the event complement. The verb *cause* is not annotated as an event.

61. *The rains caused the floods.*

62. *John caused the fire.*

9 Event Annotation in the Vaccination Corpus

For the current annotation task we will not be marking all events in a corpus, but events that are related to the vaccination debate. In order to get familiar with the vaccination debate we will first introduce the main arguments in favour and against vaccinations, which are presented in websites such as <https://vaccines.procon.org/>, where the arguments listed below can be found.

9.1 Arguments in favour and against

Among the main arguments in favour of vaccination we find the ones listed below. This is not an exhaustive list, so it might be the case that in the corpus arguments are found that are not listed here.

- Vaccines can save children's lives.
- The ingredients in vaccines are safe in the amounts used.
- Major medical organisations state that vaccines are safe.
- Adverse reactions to vaccines are extremely rare.
- Vaccines protect the "herd."
- Vaccines save children and their parents time and money.
- Vaccines protect future generations.
- Vaccines eradicated smallpox and have nearly eradicated other diseases such as polio.
- Vaccine-preventable diseases have not disappeared so vaccination is still necessary.
- Vaccines provide economic benefits for society.

Among the main arguments against vaccination we find the following:

- Vaccines can cause serious and sometimes fatal side effects.
- Vaccines contain harmful ingredients.
- The government should not intervene in personal medical choices.
- Mandatory vaccines infringe upon constitutionally protected religious freedoms.
- Vaccines can contain ingredients some people consider immoral or otherwise objectionable.
- Vaccines are unnatural, and natural immunity is more effective than vaccination.

- The pharmaceutical companies, FDA, and CDC should not be trusted to make and regulate safe vaccines.
- Diseases that vaccines target have essentially disappeared.
- Most diseases that vaccines target are relatively harmless in many cases, thus making vaccines unnecessary.

9.2 Events that should be annotated

It is not possible to provide an exhaustive list of events that are related to the vaccination debate because the goal is not to annotate events within the vaccination semantic field only, but also events that relate to opinions, beliefs, and attitudes towards vaccinations, legal actions or government campaigns that revolve around vaccinations, economic aspects of vaccinating/or not vaccinating , etc. The key question to ask is: does the event add information about the perspectives that entities (persons, institutions, companies, ...) hold about any aspect of vaccinations? Below we list several types of events that are related to the vaccination debate, but annotators should apply their own judgement when deciding whether an event has to be annotated.

- Events related to administering and testing vaccines.
- Events that refer to actions by persons and institutions in relation to vaccines.
- Events related to explaining personal experiences with any aspect of vaccines.
- Events that are mentioned when reporting discussions related to vaccines.
- Events related to disease outbreaks.
- Events that are related to how health organisations/institution, communities, groups, individuals and other entities position themselves towards vaccines.
- Events that are mentioned when discussing the personal freedom to choose in relation to vaccines. Events can refer to the choices that governments do, to personal choices, to the decisions that are made for children, to the right of parents to choose, etc.
- Events that are mentioned when discussing the relation between vaccines and religion.
- Events that are mentioned when discussing the relation between vaccines and moral attitudes (i.e, origin of vaccines ingredients).
- Events related to the achievement that vaccines have brought.
 - Vaccines save lives.
 - Vaccines protect the community.
 - Vaccines protect future generations.
- Events mentioned when talking about diseases in relation to vaccines.
 - Prevented diseases.
 - Eradicated diseases.
 - Non-eradicated diseases.
 - Diseases caused by vaccines.
 - Harmless diseases, vaccination not necessary.
- Events mentioned when talking about globalisation and disease spreading.

- Events involved in monitoring the effects of vaccines.
- Events related to the (adverse) side effects of vaccines: illnesses, symptoms, deaths, etc.
- Events related to the Immunity level provided by vaccines.
- Vaccines components, ingredients: safe or not, allergies, side effects.
- Events involved in the commercialisation of vaccines.
- Events related to legal aspects of vaccines and policy making.
- Events about the economic effect of (lack of) vaccination (less illnesses, less expenses for family and society).

9.3 Events that should not be annotated

The following events should not be annotated:

Attribution events An attribution event expresses “a relation ascribing the ownership of an attitude towards some linguistic material, i.e. the text itself, a portion of it or its semantic content, to an entity” Pareti (2015). An attribution relation is established when someone (the author of a text, someone mentioned in the text) signals an ownership relation between a third party who expresses an attitude and some text. For example, the authors of sentence “Students like annotation tasks” establishes a relation between a source (students) and a content (annotation tasks) by means of a cue (like).

Following the PARC v3.0 annotation guidelines, the text towards which an attitude is expressed can be:

- An utterance or an assertion (i.e. either reported speech, or indirect reported speech, or mixed reported speech) as in “The student said *that he would annotate*”.
- A belief, as in “the teacher thinks *that the annotation task is easy*”.
- A knowledge, as in “the task leader knows *everything about attribution*”.
- An intention, as in “the student aims *at annotating three texts per day*”. It can be the case that either an agent influences someone to perform - or not - an action, that an agent commits to performing - or not - an action, or that an agent expresses a (mental) orientation towards some state of affairs.

In the example below *alleged*, *said*, and *reported* should not be annotated as events.

63. *In an interview with Radio Südtirol Mr Holzer alleged that some chemicals in vaccines were risky, and said parents should have a free choice about child immunisations, as in Germany, Austria and Switzerland.*
64. *Trentino-Alto Adige, a mountainous Alpine region, is reported to have one of the highest vaccine refusal rates in Italy.*

For a description of attribution events, read the guidelines on Annotating Attribution.

Events that occur in sentences that provide contextual information about what is being said . In the example below, the events in the second sentence do not have to be annotated because it provides information about where the families live.

65. *A group of German-speaking parents in northern Italy are so angry about a new requirement to get their children vaccinated that they plan to seek asylum in nearby Austria. The 130 families live in Alto Adige - also known as South Tyrol - a region that was part of Austria before 1919.*

10 Annotation of specific cases

10.1 Vaccine, vaccination

Vaccination is the act of administering a vaccine, which means that it is always an event.

Vaccine is the liquid or preparation that is administered. By default it is not an event. However, it should be annotated as event when it is used in the sense of vaccination, which happens very often.

See the following examples where *vaccine* is not annotated as event because it refers to the chemical product.

66. *Mr Holzer voiced concern about Thiomersal (or Thimerosal), a mercury-based preservative used in some vaccines, and about genetically engineered vaccines.*
67. *The UK National Health Service says Thiomersal is not used in child vaccines - and adds that it poses no risk anyway.*
68. *They discovered a measles vaccine*
69. *The logistics of getting vaccines into certain areas*
70. *The routine use of mumps vaccine*

In the following example, it is not clear whether *vaccine* is used in the sense of *vaccination* , so it should not be annotated as event:

71. *Scientific studies have debunked an alleged link between vaccines and autism, as the US Centers for Disease Control point out on their website*

In the first example below, *vaccine* is annotated as event because of the construction with the support verb *get*. *Measles* should not be annotated as event because it is a modifier of *vaccine*.

72. *He got a measles vaccine*

In the next example, *vaccine* is annotated as event because we can interpret that administering vaccines is what has protected the population:

73. *Vaccines have protected many people*

10.2 Diseases

Mentions of diseases are not always annotated as events, only when they refer to cases of the illness that can be put in a timeline.

In the following example, "autism" is not an event because it does not refer to specific cases of autism, but to the illness in general.

74. *Scientific studies have debunked an alleged link between vaccines and autism, as the US Centers for Disease Control point out on their website.*

In the following examples "measles" is annotated as event because it refers to cases of measles that can be placed in a timeline, as well as *Hashimoto's disease* and autism.

75. *The highly-contagious sickness is fatal in some cases. Some other European countries, including France and Romania, have also seen more measles cases this year.*
76. *Then he got Hashimoto's disease*
77. *There were 3 cases of autism in 2016.*

10.3 Other

Mentions of the *anti-vax movement* are annotated as event because it can be placed in a timeline. It has a starting point, a development and might have an end point also.

11 Example text

Here we provide a text annotated with vaccination related events.

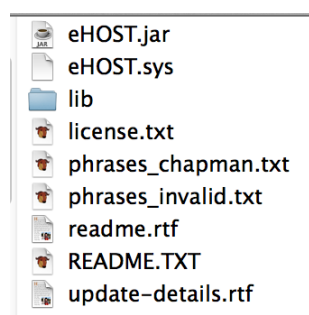
78. A group of German-speaking parents in northern Italy are so angry about a new requirement to get their children vaccinated that they plan to seek asylum in nearby Austria.
79. The 130 families live in Alto Adige - also known as South Tyrol - a region that was part of Austria before 1919.
80. Last month the Italian government ruled that children must be vaccinated against 12 common illnesses before they can enrol for state-run schools. Cases of measles have risen in Italy.
81. The highly-contagious sickness is fatal in some cases. Some other European countries, including France and Romania, have also seen more measles cases this year.
82. In some parts of Europe, including Italy, vaccination rates have dropped below those recommended by the World Health Organization (WHO).
83. The leader of the South Tyrol protest, Reinhold Holzer, said the group had sent protest messages to Italian President Sergio Mattarella, Austrian President Alexander Van der Bellen, and the UN Human Rights Council in Geneva.
84. "We won't allow our children to be poisoned. Asylum is claimed not just by people fleeing war, but also by people whose rights are being violated," said Mr Holzer, quoted by Austria's Der Standard daily.

85. Trentino-Alto Adige, a mountainous Alpine region, is reported to have one of the highest vaccine refusal rates in Italy.
86. In an interview with Radio Sdtirol Mr Holzer alleged that some chemicals in vaccines were risky, and said parents should have a free choice about child immunisations, as in Germany, Austria and Switzerland.
87. However, Germany recently announced plans to fine parents who failed to get medical advice about immunising their child.
88. Mr Holzer voiced concern about Thiomersal (or Thimerosal), a mercury-based preservative used in some vaccines, and about genetically engineered vaccines.
89. The UK National Health Service says Thiomersal is not used in child vaccines - and adds that it poses no risk anyway.
90. Conspiracy theories about the health risks of certain vaccinations - largely based on one discredited paper - have spread on the internet, prompting some parents to shun immunisation.
91. Scientific studies have debunked an alleged link between vaccines and autism, as the US Centers for Disease Control point out on their website.

12 Annotation in eHOST: Instructions for Annotators

The annotator will first read the entire text of a document and then will proceed to annotate events.

For this annotation task we use the tool eHOST.⁹ The documentation can be found at <http://ehostdoc.com/>. In order to install the tool, download the file ehost-bp-1.zip at <https://code.google.com/archive/p/ehost/downloads>, and unzip it in your preferred directory. As a result you will have a directory called ehost-bp-1 which looks like this:

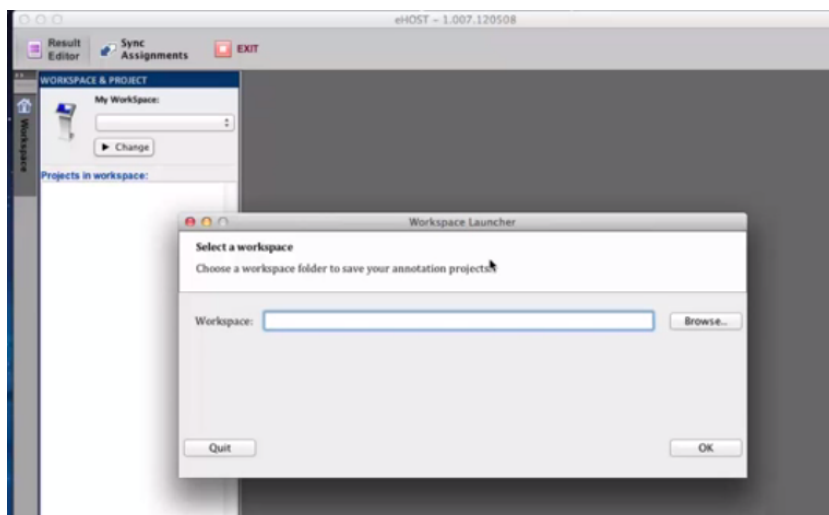


You should also have a Dropbox account, which can be made at <https://www.dropbox.com>.

12.1 Configuring the Workspace

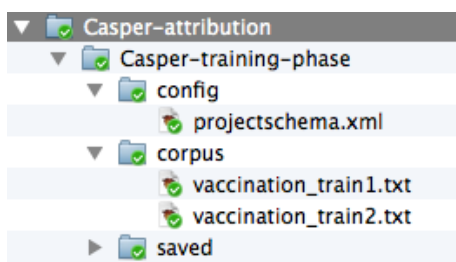
In order to launch the tool you will click in the file eHOST.jar. The first time you open eHost you will be asked to choose a workspace. A workspace is a folder with ongoing projects.

⁹<http://blulab.chpc.utah.edu/content/ehost-extensible-human-oracle-suite-tools>



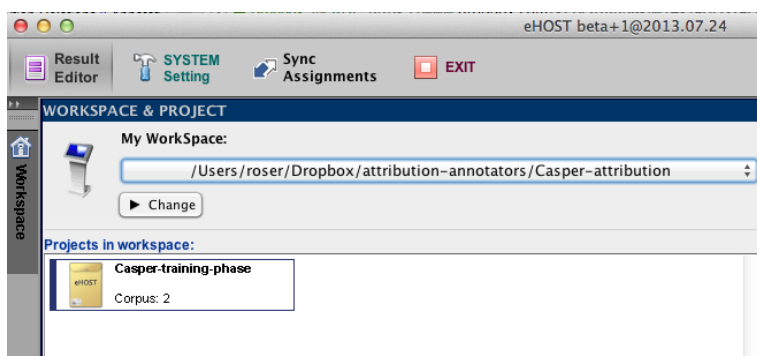
Watch now the tutorial video *Launching a workspace* to be found at <https://www.dropbox.com/s/1v3rtu8kiw14k/eHOST-Launch.mp4?dl=0>.

Configure your workspace to be the Dropbox folder that you will have been assigned by the task leader. The advantage of using Dropbox is that then your files will be stored in a safe place and the task leader can access them without you having to send them or upload them. Your workspace folder will have different projects. The first one will be one called “yourname-training-phase”. The folder should look like this:



The “config” subfolder contains the specification of the task, the “corpus” folder contains the files to be annotated, and the “saved” folder contains the saved annotations. When you start the task the “saved” folder is empty.

The selected workspace will appear in the tool interface. It will look like this:

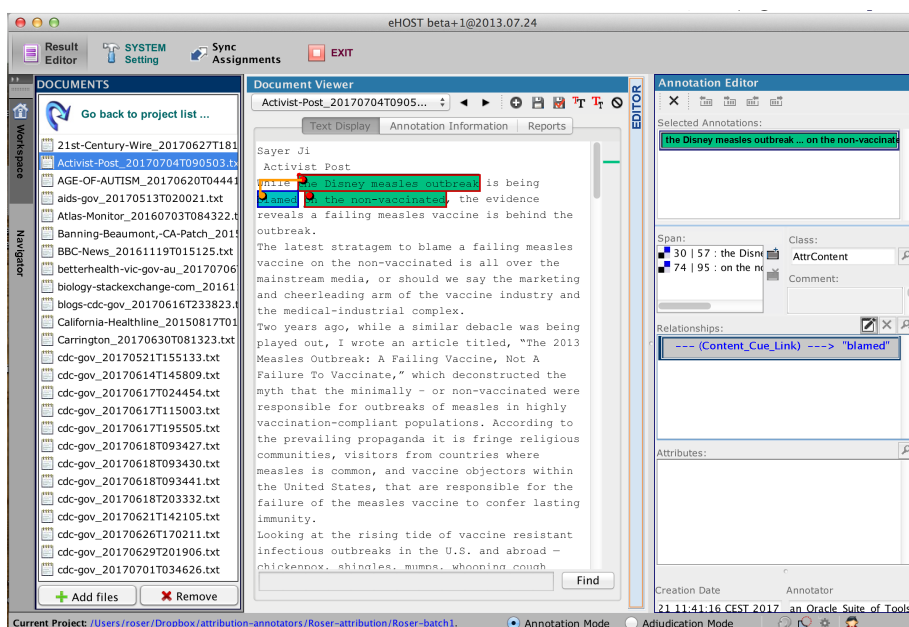


Once you have configured your workspace folder, you can start annotating. Before that, read Chapters 1 and 2 from the eHOST official Documentation Wiki (<http://ehostdoc.com/html/start.html>) and watch the following tutorial videos from the eHost website:

- *Annotating entities from a corpus to be found* at <https://www.dropbox.com/s/gafuscfb9pov48/eHOST-Annotate.mp4?dl=0>
- *Annotating relationships between entities to be found* at <https://www.dropbox.com/s/pja7jryxybfs9zg/eHOST-Relations.mp4?dl=0>
- *Saving your annotated corpus and workspace to be found* at <https://www.dropbox.com/s/tdzgrtw8uf18o3d/eHOST-Save.mp4?dl=0\%20target=>.

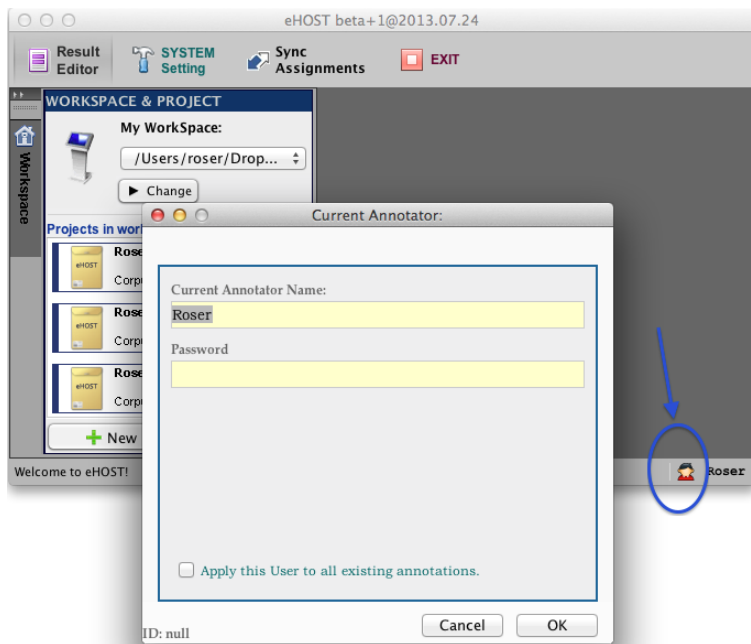
Additionally, you can watch the video *Demonstration of Annotation Tools* at <https://www.youtube.com/watch?v=T5kpPxgw9tw> from minute 35:31 to minute 58:00.

The eHost Workspace interface is divided into three panels: Documents, which contains the list of files per project; Document Viewer, which displays the text to be annotated, and Annotation Editor, where several actions can be performed for a specific markable.



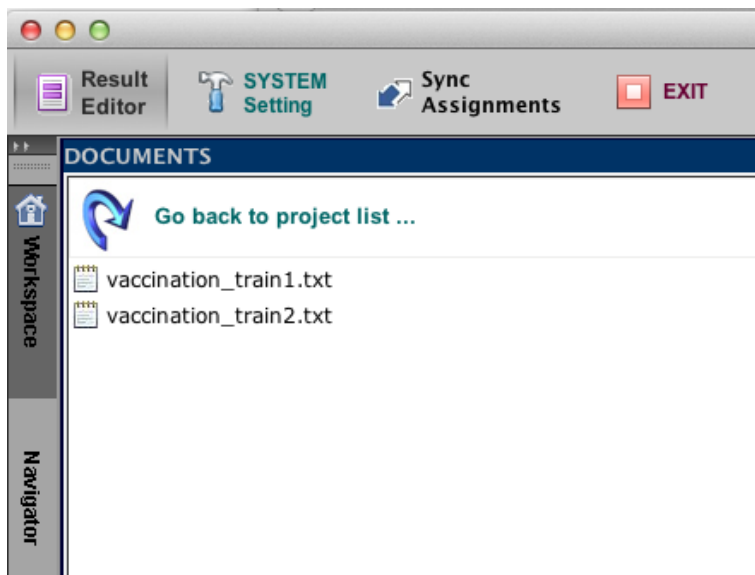
12.2 Introduce annotator name

Before you start annotating you should introduce your annotator name. You can do this by right-clicking at the person icon that appears in the right bottom corner of the interface. A pop-up window will appear where you can introduce the “Current Annotator Name”. Doing this is very important in order to be able to compare annotations by different annotators and calculate inter-annotator agreement.

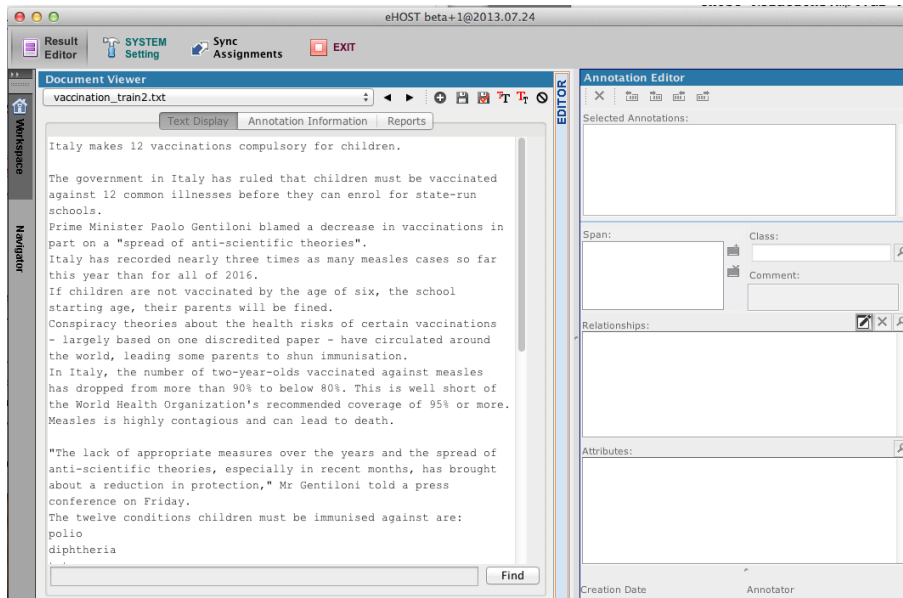


12.3 Selecting a project

In order to start the annotation task you need to select a project from the “Workspace and project” panel (see figure above). Double click on the project name. Let’s imagine that you select the project “Casper-training-phase” shown in a previous figure. You will then see a list of files in the “Documents” panel:

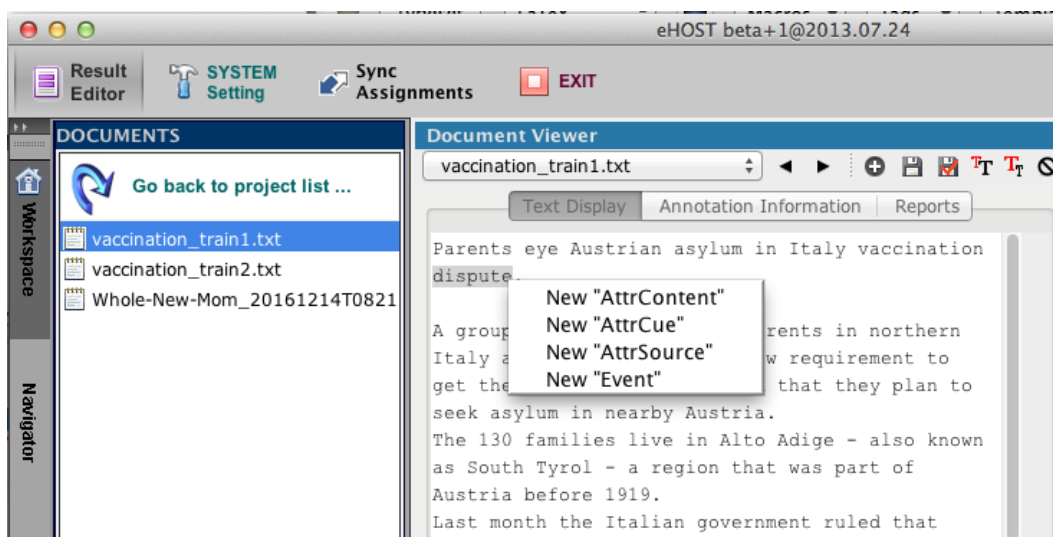


Double click in one of the files to start the annotation task. It will appear in the “Document viewer” panel. First, read the full text once. Then proceed to annotate events.



12.4 Annotating markables

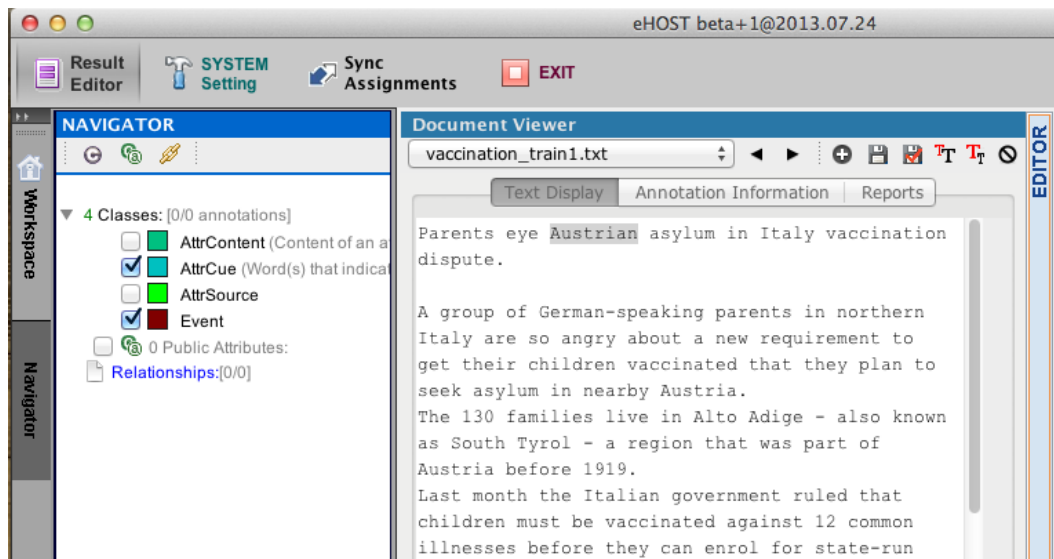
In this task you need to annotate only one markable: Event. As you have seen in the video tutorials, in order to annotate you need to be in the “Text Display” tab of the Document Viewer panel. Select the text that you want to annotate with a markable. A pop-up window will appear where you can select the markable that you want to assign.



12.5 Visualising already annotated markables

As you see, there are four markables defined for this task: AttrContent, AttrSource, AttrCue, and Event. In this annotation task you will be annotating only Events. However, it might be useful to visualise markables that have already been annotated in the text because if a piece of text has already been annotated as AttrCue, it does not have to be annotated as Event. In order to select the markables that you want to visualize, go to the Navigator Panel. There you will see the four classes defined as markables. By ticking

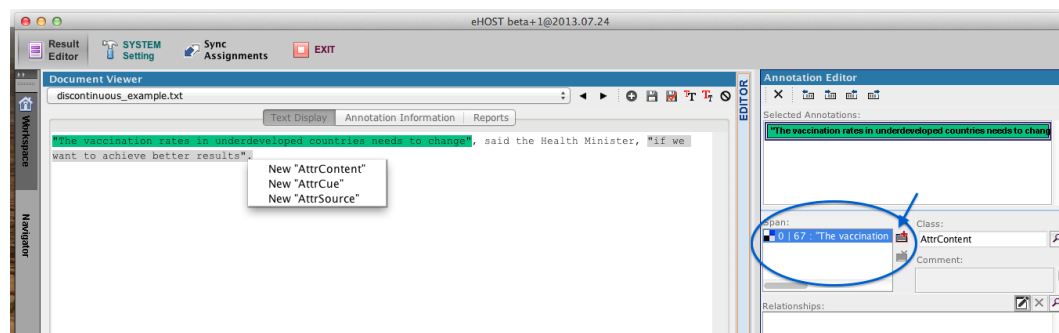
on the white box to the left of the class you can choose to visualise it. For this annotation task, please, tick on AttrCue and Event.



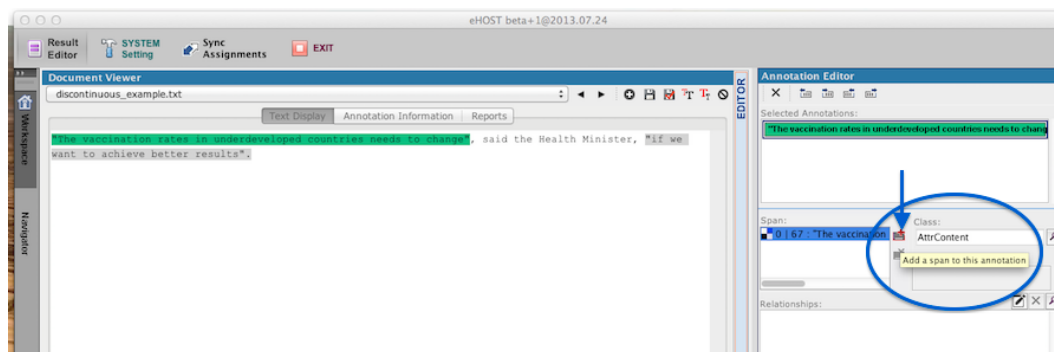
12.6 Discontinuous span for markables

It can be the case that a markable is discontinuous. In order to mark discontinuous spans of text you need to do the following:

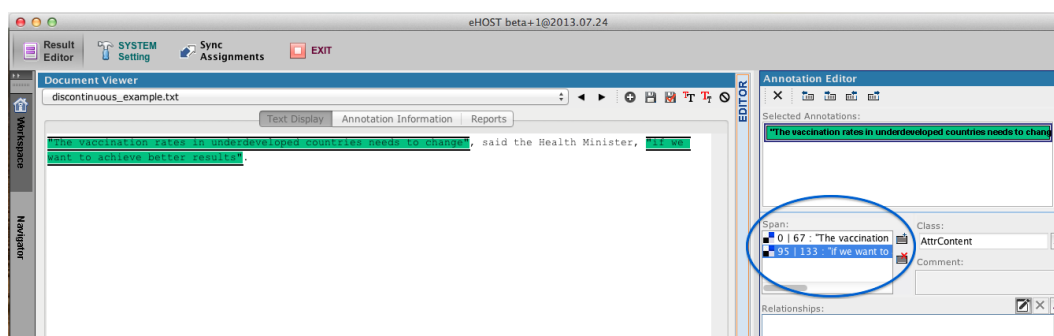
- Select the first part of the span and assign to it a class. The annotated span will appear in the Annotation Editor panel to the right of the interface.
- Then select another part of the markable span.



The menu to select a class will be displayed, but do not select a class. Go to the Annotation Editor panel, to the Span box where you will see the span of the first part of the markable that you annotated in the previous step.



To the right you will see an icon that is a folder with a + sign. Click on the icon and the span of the selected text will be added to the span of the previous step.

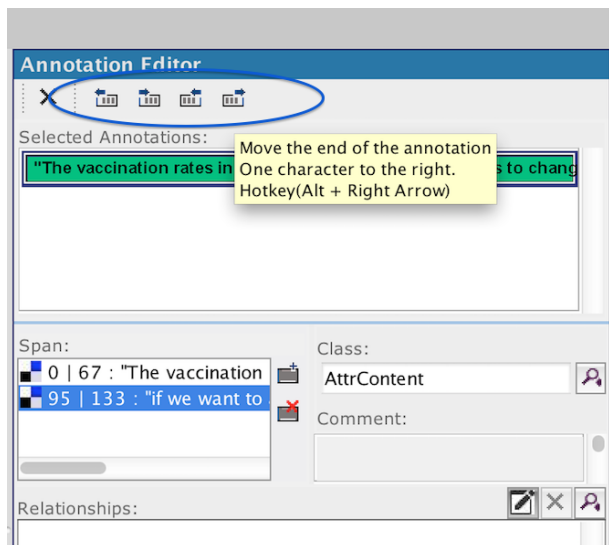


- Repeat the previous step as many times as necessary.

12.7 Adjusting the span of a markable

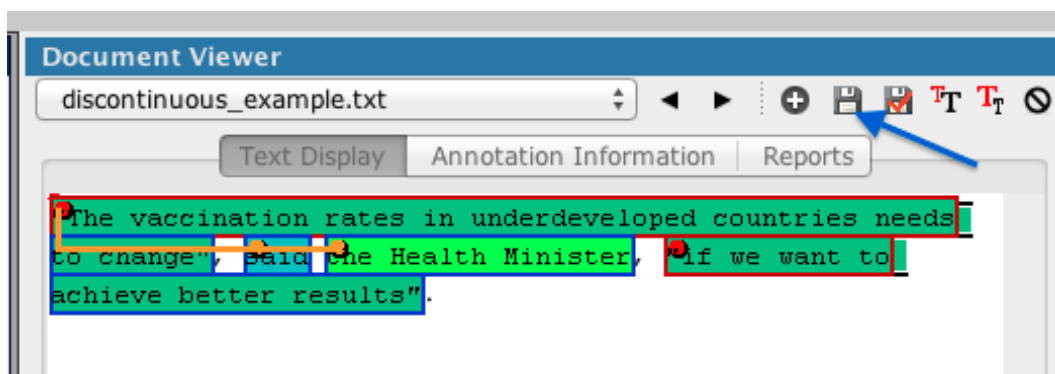
It might be the case that when you select text, the span of the markable is not correct. For example, you might want to exclude the final dot at the end of a sentence. The tool will first expand the selection to the punctuation mark attached to the word, but you can modify this by using the options in the Annotation Editor (the rectangles with arrows that appear below “Annotation Editor”).

Click on the markable whose span you want to change. Its span will appear in the Span box of the Annotation Editor. Click on the span that you want to change and press one of the 4 buttons for span modification, depending on how do you want to modify it. You can move the beginning and the end of the span one character to the right or to the left.



12.8 Saving your data

It is very important that you save your annotations frequently. eHOST does **NOT** save them for you. You can save them by clicking on the save button of the Document Viewer panel. The annotations will be saved in your Dropbox folder (assuming that you configured it to be your workspace). Thus, you do not need to send them to the task leader.



References

- Cybulska, A. and Vossen, P. (2014). Guidelines for ecb+ annotation of events and their coreference. Technical report, Technical report, Technical Report NWR-2014-1, VU University Amsterdam.
- ISO, S. W. G. (2008). *ISO DIS 24617-1: 2008 Language resource management - Semantic annotation framework - Part 1: Time and events*. ISO Central Secretariat, Geneva.
- Pareti, S. (2015). *Attribution: A Computational Approach*. PhD thesis, University of Edinburgh.
- Pustejovsky, J., Verhagen, M., Saurí, R., Littman, J., Gaizauskas, R., Katz, G., Mani, I., Knippen, R., and Setzer, A. (2006). TimeBank 1.2. Technical report, Linguistic Data Consortium (LDC), Philadelphia, PA.
- Simone, R. and Masini, F. (2007). Support nouns and verbal features: a case study from italian. *Verbum*, (1-2):143–172.
- Styler, W., Crooks, K., Hamang, M., and O’Gorman, T. (2014). *Richer Event Description (RED) Annotation Guidelines*. University of Colorado at Boulder.
- Tonelli, S., Sprugnoli, R., Speranza, M., and Minard, A.-L. (2014). NewsReader Guidelines for Annotation at Document Level. Technical report, Technical report, Technical Report NWR-2014-2, Fondazione Bruno Kessler, Trento.