



OPEN POLARITY ENHANCED NAMED ENTITY RECOGNITION



Instrument: Collaborative Project

Thematic Priority: FP7- ICT -2011.4.1

Grant Agreement: 296451

WP5 - SENTIMENT- Development of Sentiment Lexicons and Opinion Mining Techniques

Deliverable Number	WP5
Title	Guidelines for Opinion Annotations of Customer Reviews in the Tourist Domain
Version	0.1
Date	4th June 2014
Status	Draft

Project Partners: VICOMTECH-IK4, UPV/EHU, VUA, SYNTHEMA, OLERY, CNR

Contributors: VUA

Every effort has been made to ensure that all statements and information contained herein are accurate; however the Partners accept no liability for any error or omission in the same.

© Copyright in this document remains vested in the Project Partners

Document Control

Version	Date	Author	Modifications
0.1	4 June 2014	Chantal van Son	First edition.
0.1	4 June 2014	Isa Maks	First edition.



ANNOTATION GUIDELINES

English version 5.0

Table of contents

1	Introduction.....	5
2	Opinion Layer.....	6
2.1	Explanation	6
2.2	Tag levels in KAF Annotation Tool	7
2.3	Should I annotate this sentence?	8
2.3.1	Positive/negative attitude	9
2.3.2	On topic	9
2.3.3	To-the-point.....	9
2.3.4	Example	10
2.4	Opinion expressions	11
2.4.1	Strength: default vs. strong	11
2.4.2	Modifiers: intensifiers, weakeners and polarity shifters	12
2.5	Opinion Holder.....	13
2.6	Opinion Target	13
2.7	Multiple expressions, holders or targets.....	14
2.8	Nested opinions are annotated as separate simple opinions	15
2.9	2.9 Presence and absence of aspects	16
2.10	Span of words.....	16
2.10.1	Holders and targets	16
2.10.2	Opinion expressions	18
3	Aspects Layer	19
3.1	Explanation	19
3.2	Tag levels in KAF Annotation Tool	19
4	Document Level Sentiment	20
4.1	Explanation	20
4.2	Tag level in KAF Annotation Tool.....	20
5	Extra guidelines for the annotations of Attraction Reviews.....	22
5.1	Single sentence annotation approach (all languages except Dutch).....	22

5.2	Aspects and examples (Tag 3)	22
5.3	Type of attraction (Tag 5)	23
5.4	Annotating non-opinionated sentences	23

1 Introduction

The aim of this annotation project is to identify subjective evaluations regarding topics discussed in reviews – mainly hotel reviews. The annotation task that you will be performing consists of two subtasks, in which you will identify:

- (1) Opinions frames (an opinion from someone about something)
- (2) Aspects (of the entity being reviewed)

In the Quick Start Guide the very basis of the annotation of opinions and aspects in the KAF Annotation Tool are explained, which can be used to get a basic idea of the task and to quickly start with the annotation. The Quick Start Guide also provides an explanation of the basics of using the KAF Annotation Tool as well as a step-by-step demonstration of the annotation of a few simple sentences in the tool.

These Annotation Guidelines provide more elaborate information about how to decide on the words to annotate and the correct labels to assign, with a number of illustrative examples and discussions on the problems you might encounter while annotating. Here we only deal with the theoretical issues of the annotation and do not explain how to use the KAF Annotation Tool. Therefore, we advise you to first read the Quick Start Guide for learning the basics and how to work with the KAF Annotation Tool; the Annotation Guidelines can be read in addition for more detailed information about ‘making the right decisions’.

2 Opinion Layer

2.1 Explanation

In this layer, opinion triples are identified. A complete opinion triple consists of the following three components, with the opinion expression as the core element:

- I. Opinion expression
 - Negative
 - Positive
 - Strong negative
 - Strong positive
- II. Opinion holder - whose opinion is expressed (optional)
- III. Opinion target - what is the opinion about (optional)

The opinion expressions are further subcategorized according to their orientation (positive vs. negative) and according to their strength (default vs. strong), resulting in 4 different labels in the KAF Annotation Tool. How to decide on these labels will be discussed in Section 2.4.

Opinion targets and opinion holders are only identified if they occur in the same sentence as the opinion expression. An example of a complete opinion triple is sentence (1), in which *I* is the opinion holder, *avoid* the opinion expression, and *this hotel* the opinion target. We will use the colors green, blue and red to mark the holder, target and expression respectively for all the examples throughout these Annotation Guidelines.

- (1) I would avoid this hotel.

However, complete opinion triples like this one rarely occur in actual reviews: in many cases the opinion holder and/or the opinion target occur in another sentence than the opinion expression, or are not explicitly lexicalized at all. For example, in (2) the opinion consists only of the opinion expression and the opinion target. In (3) there is no opinion target, but there is an opinion holder and an opinion expression. In (4) it might be possible to interpret *we* and *the hotel* in the first sentence as the holder and the target of the expression *fantastic* in the second sentence, but since they do not occur in the same sentence, they are not annotated as such. These issues will be further discussed in Sections 2.5 and 2.6.

- (2) The rooms were nicely decorated.
(3) We had a great time!
(4) We arrived rather early in the hotel. Fantastic!!

Furthermore, there may be multiple holders, targets or expressions in one opinion. In (5), there are two opinion holders: *my boyfriend* and *I* (note that there is only one opinion expression, *did not like at all*, which is split in half by the opinion target). In (6) there are two targets, *the owner* and *his staff* and (7) contains two opinion expressions, *small* and *had a great view*. This will be further discussed in Section 2.7.

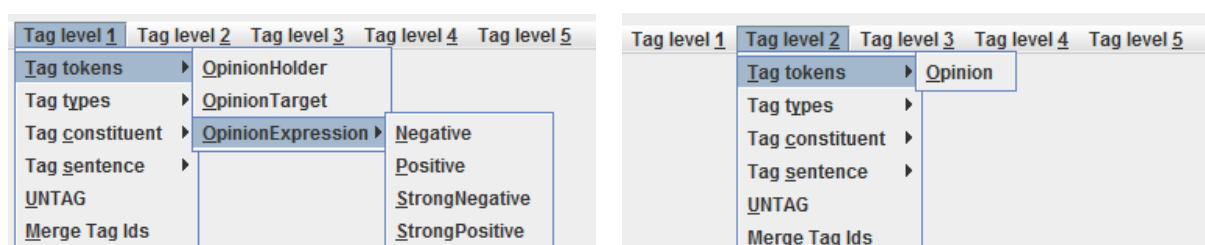
- (5) *My boyfriend* and *I* *did not like this hotel at all*.
 (6) *The owner* and *his staff* were *very friendly*.
 (7) *The balcony* was *small* but *had a great view*.

Finally, opinions can be nested, as opinion (8b) is nested within the ‘higher-level’ opinion of (8a). This will be discussed in Section 2.8.

- (8) a. *Average hotel with poor Internet Service*.
 b. Average hotel with [*poor Internet Service*].

2.2 Tag levels in KAF Annotation Tool

There are four tag levels in the KAF Annotation Tool that are used to annotate the opinion frames: Tag levels 1-4. Tag level 1 is used for the tagging of the three components of an opinion triple; the tag labels of this level are represented in Table 1. Tag level 2 contains just one label ‘Opinion’ and is used for linking the separate components of an opinion together.



For tagging sentence (1), repeated here in (8), you will thus use both Tag level 1 and Tag level 2. First annotate the separate components of the opinion in Tag level 1 (*I* as opinion holder, *avoid* as negative opinion expression and *this hotel* as opinion target). Then select the span of words that form the opinion as a whole (here, the whole sentence) and use the tag ‘Opinion’ from Tag level 2 to link them together.

- (9) *I* would *avoid this hotel*.

Tag level 3 and 4 are a repetition of Tag level 1 and 2 and contain the same labels; these levels are only used in some cases with multiple holders, target and/or expressions and for the annotation of nested opinions. How to use these tag levels will be explained in Sections 2.7 and 2.8.

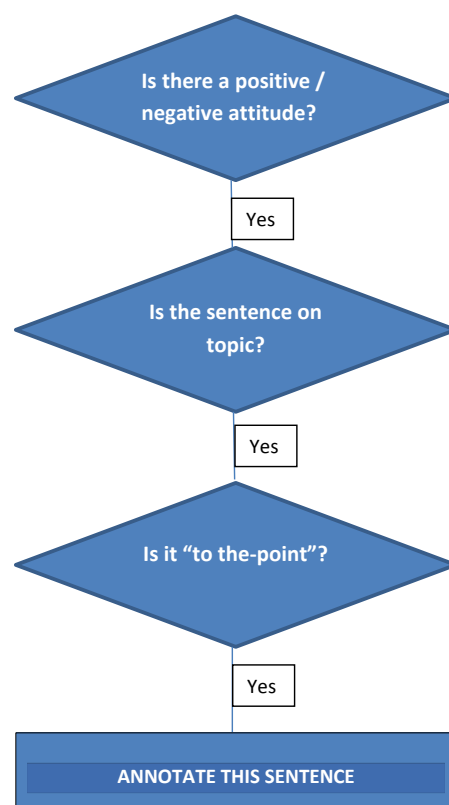
	Label	Explanation
1	OpinionHolder	word or span of words that refer to the opinion holder (e.g. <i>I, we, my boyfriend</i>)
2	OpinionTarget	word or span of words that refer to the opinion target (e.g. <i>the hotel, our room, the view</i>)
3	OpinionExpression - Negative	the span of words that refer to the opinion expression with negative polarity (e.g. <i>bad, unfriendly, poppycock, not very good</i>)
4	OpinionExpression - Positive	the span of words that refer to the opinion expression with positive polarity (e.g. <i>nice, still worth it to check out, etc.</i>)
5	OpinionExpression - StrongNegative	the span of words that refer to the opinion expression with strong negative polarity (e.g. <i>extremely noisy, horrible, a disaster, etc.</i>)
6	OpinionExpression - StrongPositive	the span of words that refer to the opinion expression with strong positive polarity (e.g. <i>excellent, the same wow factor, the best thing</i>)

Table 1: Explanation of the tag labels of Tag level 1 and 3

2.3 Should I annotate this sentence?

The first step of the annotation task in the Opinion Layer is to ask yourself: should I annotate this sentence (or part of it) or not? The basic idea is that not all the review text is opinionated and/or contains information that should be included in the annotation process. Therefore, annotators should ask themselves - for each sentence- the following questions. Only if all questions are positively answered, the sentence will be taken into consideration for further annotations.

The next Sections will give some examples of (parts of) sentences that should not be annotated, because they lead to a negative answer for one or more of these questions.



2.3.1 Positive/negative attitude

Does the sentence or part of a sentence express an attitude (either positive or negative)?

The following sentences contain no opinion or judgment and are therefore excluded from further annotations. The questions may also affect parts of sentences like subordinating clauses (e.g. in sentence (7) in Section 2.3.2).

- (10) Me and my wife stayed at the English inn in July 2012 and can't remember the weekend but we had A.C.
- (11) I found this place solely based on a last minute online research.
- (12) The ratings I stumbled upon gave this hotel a rather "best of Charlottesville" recommendation.

2.3.2 On topic

Is the sentence “on-topic”? The topic of the reviews is the touristic attraction, restaurant or hotel itself. It may have many aspects which are all considered “on-topic” including its surroundings, nearby attractions. Sometimes reviewers give opinionated information which seem to refer to other aspects of their stay (or life) than the topic under evaluation. The following sentences are considered not “on-topic”.

- (13) We were driven by a friendly taxi driver, who'd picked us up at the ferry port.
- (14) We arose to a beautifully sunny day and headed down for our breakfast.
- (15) Considering that the other Athenians I met are not so polite or friendly (they do have a financial crisis going on so this is fully understandable) ...
- (16) We also did a night time 2 and a half hour canal cruise including 4 course dinner & drinks - it was great... And a trip around the city by horse drawn carriage which was worth every penny that was for 2 hours as well.

However, keep in mind that there may be opinions that *indirectly* refer to the entity being reviewed. For example, in (17) a positive thing is said about other hotels than the hotel that is being reviewed. But because this opinion should be interpreted as a comparison with the hotel under evaluation, it does say something about the latter as well. Therefore, these kind of opinions are considered on-topic and should be annotated.

- (17) There are other hotels with more attentive, friendly service.

2.3.3 To-the-point

Is it “to-the-point”? A reviewer gives his opinion and then elaborates with lots of details on it. These “stories” are opinionated and on-topic but they elaborate a lot on things already said. In the following fragments, the underlined sentences are considered not to-the-point:

- (18) Our room was very small; we could hardly get past the furnishings and into the bathroom. Indeed, the latter was almost as big as the bedroom.
- (19) Then I looked over at the bed and saw it. A dead mouse. Right there in the middle. Mice are bad enough, but how did this thing just die in the middle of the bed? I looked at it for a while, gathered my stuff and went downstairs. I showed a picture of it to the staff. I told them that changing the room would not be sufficient, you understand.
- (20) The bathroom was better in that it felt a bit more modern with a jacuzzi bath but aspects of it were poorly designed such as the towel rail inside the bath enclosure made it difficult to access the towels easily.

This is also an important point for the span of words of opinion expressions (Section 2.10.2).

2.3.4 Example

In the example below, which can also be found in the example folder of the KAF Annotation Tool (Electra1), the bold fragments are considered opinionated and are taken into consideration for further annotations in the Opinion Layer. The next Section will explain how to annotate these sentences. The non-bold fragments are skipped as they are considered “non-topical”, “not to-the-point” or “not-opinionated”.

“Not a 5* hotel”

Reviewed January 30 , 2013 via mobile

I visit Athens twice a year for business but this was my first stay at the Electra Palace. **The lobby is impressive, but my room did not have an acropolis view and was disappointing. The finishes are tired, with cheap fake wood floors and old-fashioned decor. There was a small balcony but with no access and it also suffered from traffic noise from the street** (room on 2nd floor) and **thin walls so noise from other guests. It lacks practical touches,** like sockets by the bed to charge your phone, tea/coffee-making facilities and **wifi which is available in the room but at a high price** (20 euros for 3 days). **I would not stay here again - there are much nicer hotels with lower but more accurate star ratings in the vicinity, at a more reasonable rate, with more modern facilities and more attentive, friendly service.**

2.4 Opinion expressions

The opinion expressions are subcategorized according to their orientation (positive vs. negative) and according to their strength (default vs. strong). The following paragraphs will explain how to decide whether the opinion has a positive or a negative polarity, whether the opinion is strong or ‘normal’, and how both polarity and strength of the expression can be modified by *intensifiers*, *weakeners* and *polarity shifters*.

2.4.1 Polarity: positive vs. negative

The core element of an opinion expression is a single word, a phrase or a clause that carries a meaning with positive or negative polarity, like *good* or *still worth it to check out*. They include evaluations, appreciations, judgments of behavior (24) and also more emotional and attitudinal stances like in (26) and (27). When deciding whether a word (or span of words) is opinionated and whether this is positive or negative, the annotators should try to understand what is meant by the reviewer instead of following their own preferences. For example, even when the annotator happens to like small rooms, it is clear from the context “we could hardly get past the furnishings and into the bathroom” that the reviewer who wrote sentence (22) does not. In this case negative polarity is annotated (strong negative polarity in this case; see also Section 2.4.3 about strength).

- (21) There is a **poor** Internet service (-)
- (22) **Our room** was very small; we could hardly get past the furnishings and into the bathroom. (--)
- (23) **The area** has the feel of a neighbourhood (+)
- (24) **The staff** was ungracious (-)
- (25) **The rooms** were very clean (++)
- (26) **We** enjoyed the view from the rooms (+)
- (27) **I** would avoid this hotel (-)
- (28) And **the room** lacked atmosphere (-)

2.4.1 Strength: default vs. strong

Some opinion words, phrases or clauses express stronger polarity than others, like *luxurious* (vs. *good*) and *horrible* (vs. *bad*). In these cases, the labels “strong positive expression” or “strong negative expression” are used. Also the use of capitals, repetition and exclamation marks may affect the strength of the opinion word. For example, the expression in (31) is strong because of the use of the exclamation marks and capitals. The expression in (35) is strong because of the repetition.

- (29) **The bed** and **bedding** were luxurious (++)
- (30) **The maid service** were excellent (++)
- (31) **BAD!!!!** (--)
- (32) Look, **it** seriously ruined our trip to Athens (--)

- (33) **The olive shampoo** and **body gel** was top grade - I took an entire set each of three days that will supply me at home for a month. (++)
- (34) **The lobby** is impressive (++)
- (35) Lies. Lies. Avoid Avoid. (--)

2.4.2 Modifiers: intensifiers, weakeners and polarity shifters

The core element of an opinion expression, i.e. the words that carry a positive or negative meaning, may also be combined with modifiers. These modifiers do not express polarity themselves, but change the polarity or strength of the words to which they refer. The core elements together with these modifiers function as a unit and are annotated as a single opinion expression. There are three types of modifiers: *intensifiers*, *weakeners* and *polarity shifters*.

Intensifiers are words or spans of words that intensify the strength of an opinion word. If the opinion word that is being modified is positive, the resulting polarity of the whole opinion expression is “strong positive”. If the opinion word that is being modified is negative, the resulting polarity of the whole opinion expression is “strong negative”. In the following examples, the intensifiers are underlined.

- (36) **Our room** was very small (--)
- (37) This hotel employs the most friendly efficient **staff** you could wish for (++)
- (38) The biggest penny pinching **I have ever heard of** (++)
- (39) Look, it seriously ruined our trip to Athens (--)
- (40) Too much noise (--)
- (41) **The room** was very beautiful (++)
- (42) I managed to squeeze the rental car down the insanely tight and steep **driveway** into the garage (--)

Weakeners are words or spans of words that weaken the strength of an opinion word. Since we do not distinguish between ‘normal’ and ‘weak’ polarity, the resulting polarity is always ‘normal’ (default). For example:

- (43) **The standard room** was relatively large (-)
- (44) **I** only have a small complaint (-)
- (45) **The bath room** was a bit more modern (+)

Polarity shifters are words or spans of words that shift the polarity of an opinion word from positive to negative or vice versa, like negations or *too*. For example, in (47) the positive polarity of the word *friendly* is shifted by the polarity shifter *too*, resulting in a negative opinion expression.

- (46) **The hotel's location** was not a very interesting part of town (-)
- (47) **I** have no complaint whatsoever (+)

- (48) Will never return (-)
- (49) They are too friendly
- (50) My room did not have an acropolisview (-)
- (51) .. but with no access .. (-)
- (52) I would not stay here again (-)

2.5 Opinion Holder

In reviews, the opinion holder is only in few cases explicitly mentioned. Rather, he or she is *implicitly* present as the writer of the review. For example, in sentences (53-55), it is clear that the opinions are held by the writer of the review, but none of them contain an explicit reference to him or her. In these cases no opinion holder is annotated. In example (56) the writer of the review is lexicalized with *I*; this is annotated as the opinion holder. Example (57) shows that the opinion holder can also be another person than the writer of the review. As already said, an opinion holder must occur in the same sentence as the opinion expression. For example, example (58) has an opinion expression (*wonderful*) in the second sentence, but *we* of the first sentence is not labeled as opinion holder because these two elements occur in different sentences.

- (53) The staff was ungracious. [no opinion holder]
- (54) The rooms were very clean. [no opinion holder]
- (55) And the room lacked atmosphere. [no opinion holder]
- (56) I think this is the most beautiful temple of the world.
- (57) Other guests were complaining about the rooms as well.
- (58) We went to Switzerland. That is a wonderful country. [no opinion holder]

2.6 Opinion Target

Opinion targets refer to the topic of the opinion, to what the opinion is about. In reviews, there are ‘supertargets’ and ‘subtargets’, which are both annotated as opinion target. The ‘supertarget’ is the entity evaluated by the reviewer (hotel, product, attraction, etc.). Sometimes it is explicitly mentioned and annotated as target, as in (59). Usually, however, the supertarget is implicit, as in (60). In that case the opinion expression is annotated, but the target is not explicitly present, so this cannot be annotated. The opinion expressions will then be interpreted as having the ‘default’ supertarget.

- (59) This is a beautiful hotel
- (60) Fantastic!! [no target]

The target of the opinion may also refer to a ‘subtarget’, i.e. an aspect of the supertarget (see also Section 3 for the annotation of aspects). For example, a hotel may have several subtargets like *the rooms* in (61), *the view* in (62), *the location* in (63), *Internet Service* in

(64) and *the staff* in (65). If these are mentioned in the text, they are annotated as the opinion target of that expression.

- (61) **The rooms** are beautiful.
- (62) **The view** is fantastic.
- (63) **The location** was superb.
- (64) One of the most annoying things is the poor **Internet Service**.
- (65) **The staff** was ungracious.

Again, the target must be in the same sentence as the opinion expression. So in the second sentence of (66), no target is annotated (since *the hotel* occurs in another sentence). But: targets may also be pronouns that refer to entities. So in (67) and (69), *it* refers to (the location of) the hotel. These pronouns are annotated as targets as well.

- (66) We arrived rather early in the hotel. Fantastic. [no target]
- (67) We arrived rather early in the hotel. **It** is fantastic. [no target]
- (68) The location of this hotel was perfect. **It** was walking distance to markets and restaurants as well as the Acropolis.

2.7 Multiple expressions, holders or targets

An opinion may consist of multiple expressions, holders or targets. For example, in (69) there are two opinion targets; in (70) there are three opinion expressions, and in (71) there are two opinion holders.

- (69) At least **the bathroom**¹ and **bed**² are perfectly clean. (Figure 1)
- (70) **The bathroom** is **simple**¹ but **clean**² and **sufficient**³. (Figure 2)
- (71) **Me**¹ and **my wife** really enjoyed our stay **here**!

In all of these cases, these components are all linked together with the “Opinion” tag from Tag level 2. However, make sure that in Tag level 1 the components get their own tag id. So for example, in (69) *the bathroom* is separately annotated from *bed* as two opinion targets, giving them both their own tag id at Tag level 1. However, they are part of the same opinion, so they have the same tag id at Tag level 2 (Figure 1). In (70), all three opinion expressions get their own tag id, but are annotated as one opinion (Figure 2).

Id	Word token	Pos	Tag1	Tag1 id	Tag2	Tag2 id
w_66	.			0		0
w_67	At			0		0
w_68	least	N		0		0
w_69	the		OpinionTarget	1	Opinion	5
w_70	bathroom		OpinionTarget	1	Opinion	5
w_71	and			0	Opinion	5
w_72	bed		OpinionTarget	2	Opinion	5
w_73	are			0	Opinion	5
w_74	perfectly	A	StrongPositive	4	Opinion	5
w_75	clean	N	StrongPositive	4	Opinion	5

Figure 1: Example of opinion with two opinion targets

Id	Word token	Pos	Tag1	Tag1 id	Tag2	Tag2 id
w_63	!!			0		0
w_64	The		OpinionTarget	1	Opinion	5
w_65	bathroom		OpinionTarget	1	Opinion	5
w_66	is			0	Opinion	5
w_67	simple		Negative	2	Opinion	5
w_68	but			0	Opinion	5
w_69	clean	N	Positive	3	Opinion	5
w_70	and			0	Opinion	5
w_71	sufficient	G	StrongPositive	4	Opinion	5
w_72	.			0		0

Figure 2: Example of opinion with three opinion expressions

2.8 Nested opinions are annotated as separate simple opinions

Opinions can be nested, that is, one opinion may be expressed *within* another opinion. Sentence (73) is an example of such a nested opinion. This sentence contains one opinion with the target *room* and 5 opinion expression that refer to this target: *clean*¹, *modern*², *comfortable*³, *with comfy bed*⁴ (as represented in 73a). However, within this opinion, we can see another opinion target *bed* with the opinion expression *comfy* (as represented in 73b).

- (73) a. Clean, modern, comfortable room with comfy bed.
b. Clean, modern, comfortable room with comfy bed.
c. Clean, modern, comfortable room with comfy bed.

In (73), *bed* is both part of the positive expression in the ‘higher-level’ opinion, whereas it functions as the opinion target in the embedded opinion. *Comfy* is in the higher-level opinion part of the broader opinion expression *with comfy bed*, but in the nested opinion it is an independent expression.

Although in cases like this one, some words serve several functions, only one of these functions is annotated. In this case we consider *clean, comfortable room* as one expression and *comfy bed* as another one (cf. *second one*).

2.9 Presence and absence of aspects

Regularly just the presence or absence of an aspect of the entity being reviewed is regarded as positive or negative by the reviewer. For example, the presence of *a swimming pool* and *a printer* in (75) and (76) is presented as a positive thing. On the other hand, the absence of an *elevator* and *a kettle for that early morning coffee* in (77) and (78) is presented as something negative. However, the presence of something does not always have to be positive, as illustrated by (79), and the absence of an aspect is not necessarily a negative thing, as illustrated by (80).

(74) -

(75) **There is a swimming pool.** (+)

(76) **A printer is available for guests use.** (+)

(77) **It has no elevator.** (-)

(78) **No kettle for that early morning coffee.** (-)

(79) **There were mice and other pest problems.** (-)

(80) **There were no bedbugs.** (+)

In these cases, the aspect that is present or absent is annotated as the opinion target, and the phrase that indicates the presence or absence (e.g. *there was (no), it has (no), (no) ... available* etc.) is annotated as either a positive or a negative opinion expression. This way also other opinion expressions that might follow regarding this present or absent target can be linked to the target. For example, in (81) the presence of the opinion target *a balcony*, expressed by the phrase *there was*, is regarded as a positive thing; however, this target is subsequently negatively opinionated with the expression *really small*.

(81) **There was a balcony, but really small.** (+) / (-)

2.10 Span of words

It sometimes may be difficult to decide on the exact span of words to annotate for holders, targets and expressions. Reviewers often provide additional information about the entity that they refer to or why they think something was good or bad. The main point to keep in mind is: try to be as minimal as possible, so only include the words that directly refer to the entity or judgment (that is, the words that are necessary to understand the key point) and avoid including any supplementary information or vague descriptions.

2.10.1 Holders and targets

For opinion holders and opinion targets, only the words that directly refer to the entity are part of the span of words. This is usually the pronoun or the noun or verb that is the head of the phrase, but it may also include determiners and adjective phrases, like:

- **Articles:** *a, an, the*
- **Demonstratives:** *this, that, these, those*
- **Possessives:** *my, your, his, her, its, our, their*
- **Quantifiers:** *a few, all, another, little, many, several, both, most (of the), some (of the), etc.*
- **Adjective phrases:** *young, old, red, wooden, (the owner) of the hotel, (the rooms) on the second floor, etc.*

For example, the full phrases of the next sentences are annotated as the opinion targets:

- (82) We loved **this hotel**.
- (83) **The welcome staff** is very nice.
- (84) **Our Quaint room** had exposed beams, sloping roof and large windows.
- (85) **Some of the rooms** are tiny.

However, adjectives will often express some positive or negative evaluation and are therefore in many cases not part of the opinion target, but function as a separate opinion expression, like *beautiful* in (86) is a positive expression about the target *hotel*.

- (86) This **beautiful hotel** is a great place to stay.

In this case you should not annotate the complete phrase as a target, but only *hotel* can be the target and *beautiful* is the opinion expression (and *a great place to stay* is the second opinion expression referring to the target *hotel*).

Do not include expressions that merely provide supplementary information that is not needed to understand which entity is referred to, like appositions or non-restrictive relative clauses. These kind of expressions are often set off by comma's, like the underlined parts in (87), (89) and (91). These parts of the sentence are not annotated. Only the red/green words are annotated as part of the opinion target/holder. Also note the differences between these sentences and (88) and (90), in which the adjectival phrases are needed to know which entity is referred to and are thus included as part of the targets.

- (87) **Our room**, which was on the second floor next to the elevator, was **really small**.
- (88) **The rooms on the second floor** were **certainly the best ones available**.
- (89) **The owner of the hotel**, an old man with apparently no teeth, seemed **very unfriendly**.
- (90) **The older staff** seemed **very unfriendly**. [as opposed to the young staff]

- (91) **My boyfriend**, who had never been to Greece before, was **very pleased** with the **friendly people at our hotel**.

2.10.2 Opinion expressions

For opinion expressions it is important to limit yourself to the words that express some positive or negative meaning (optionally in combination with modifiers) and are “to-the-point” as much as possible (see also Section 2.3.3). Do not include comprehensive elaborations on the ‘why’ of some opinion, or numerous (descriptive) examples used by the reviewer to illustrate his or her point. In the following examples, only the words between brackets are considered part of the opinion expression; the underlined words are not annotated.

- (92) **This hotel** is **well located** if you are in Amsterdam for business and go to the RAI center.
- (93) **The shower** was **broken**, took me about 20 minutes to get it working!
- (94) It became **quite noisy** at odd hours of the morning or night when the tables and chairs were being set up and moved around, or when there were a lot of people walking around above.
- (95) **The bed (mattress)** was **not comfortable at all**, after spending 3 nights in NH Carlton, every morning I woke up with the backache.
- (96) I had **some issues** with **the lights in the bathroom**, which were blinking so much I had to switch them off.
- (97) **The bathroom** was **not good**, when you take a shower all the place get wet!
- (98) **The stair** was **very steep** which made it hard to carry luggage up.

3 Aspects Layer

3.1 Explanation

Opinions refer to the entity that is reviewed (e.g. hotel, touristic attraction, restaurant, product, etc.) or to an *aspect* of this entity. Aspects may be either:

- *components* or *subparts* of the entity under review (e.g. rooms and bathrooms);
- *properties* (e.g. cleanliness, value-for-money and location), or;
- other concepts *related* to the hotel (e.g. staff and all kinds of facilities).

These aspects are annotated by identifying and labeling the words that refer to them.

3.2 Tag levels in KAF Annotation Tool

In the KAF Annotation tool, Tag level 3 is used for tagging the aspects. The defined set of aspects for hotel reviews is represented in Table 2; for reviews of other entities new aspects need to be added.

Aspect Labels	Example words indicating aspects
Bathroom	bath, hair dryer, mirror
Beds	sleep, blanket, mattress
Breakfast	egg, breakfast, coffee, tea, bread
Car parking	parking, garage
Cleanliness	clean, blood, stain, dirty
Facilities	tv, airconditioning
Interior/exterior	lobby, garden, entrance, walls, old-fashioned
Internet	wifi, internet
Location	walking distance, central station, in the city center, traffic noise
Reservation/check-out/check-in	check-in, check-out
Restaurant	dinner, eat, food, delicious, specialties
Rooms	view, balcony
Staff	service, manager, kind, attentive, friendly
Swimming pool	swim, pool
Transportation	tram, public transport, metro
Value-for-money	reasonable rate, expensive

Table 2: The tag labels of Tag level 5

4 Document Level Sentiment

4.1 Explanation

Although the review contains different opinions, it also expresses one overall opinion towards the hotel. This opinion can be placed on a scale from 1 to 5 (where 1 is highly negative and 5 is highly positive). The scale refers to an answer at the following question:

- a. Did the reviewer like the hotel ? (-1 = not at all; 5 = absolutely)

4.2 Tag level in KAF Annotation Tool

In the KAF Annotation tool, Tag level 4 , first position (cf. figure 3) is used for tagging the document level sentiment. aspects. The defined set of values for document level sentiment is represented in the following table.

Value Labels	Explanation
1	highly negative
2	very negative
3	negative
4	slightly negative
5	average-negative
6	average-positive
7	slightly positive
8	positive
9	very positive
10	highly positive

Figure 3: Example of document level sentiment (tag 4)

KAF Annotation Tool										
File Search Other Tag level 1 Tag level 2 Tag level 3 Tag level 4 Tag level 5										
Id	Word token	Pos	Tag1	Tag1 id	Tag2	Tag2 id	Tag3	Tag3 id	Tag4	Tag4 id
w_1	Not	A	Negative	77	Opinion	78		0		
w_2	a	D	Negative	77	Opinion	78		0		
w_3	5*	O	Negative	77	Opinion	78		0		
w_4	hotel	N	Negative	77	Opinion	78		0		
w_5	Reviewed	V		0		0		0		
w_6	January	R		0		0		0		
w_7	30	O		0		0		0		
w_8	.	O		0		0		0		
w_9	2013	O		0		0		0		
w_10	via	P		0		0		0		
w_11	mobile	G		0		0		0		
w_12	I	N		0		0		0		
w_13	visit	N		0		0		0		
w_14	Athens	R		0		0		0		
w_15	twice	A		0		0		0		
w_16	a	D		0		0		0		
w_17	year	N		0		0		0		
w_18	for	P		0		0		0		
w_19	business	N		0		0		0		
w_20	but	C		0		0		0		
w_21	this	D		0		0		0		
w_22	was	V		0		0		0		
w_23	my	Q		0		0		0		
w_24	first	G		0		0		0		
w_25	stay	N		0		0		0		
w_26	at	P		0		0		0		
w_27	the	D		0		0		0		
w_28	Electra	R		0		0		0		
w_29	Palace	R		0		0		0		
w_30	.	O		0		0		0		
w_31	The	D	OpinionTarget	82	Opinion	76		0		
w_32	lobby	N	OpinionTarget	82	Opinion	76	Interior/exterior	163		
w_33	is	V		0	Opinion	76		0		
w_34	impressive	G	StrongPositive	80	Opinion	76		0		
w_35	.	O		0		0		0		
w_36	but	C		0		0		0		
Text: Not a 5* hotel Reviewed January 30 , 2013 via mobile I visit Athens twice a year for business but this was my first stay at the Electra Palace . The lobby is impressive , but my room did not have an acropolis view and was disappointing , was a small balcony but with no access and it also suffered from traffic noise from the street (room on 2nd floor) and thin walls so noise from other guests . It lacks practical touches , like sockets by the bed to charge your phone , tea/ 20 euros for 3 days) . I would not stay here again - there are much nicer hotels with lower but more accurate star ratings in the vicinity , at a more reasonable rate , with more modern facilities and more attentive , friendly service .										

5 Extra guidelines for the annotations of Attraction Reviews

These guidelines can be used in combination with the OpeNER guidelines for opinion annotations of hotel reviews. With respect to the guidelines of the hotel reviews there are 2 modifications:

- * Extra aspects (tag 3) have been added to cover the new topics
- * An extra tag level (tag 5) has been added to indicate which kind of attraction or review is described

5.1 Single sentence annotation approach (all languages except Dutch)

We follow the single sentence annotation approach for all languages except Dutch. So, we are **not** going to annotate complete reviews, but only one sentence from each review. The idea is to annotate the first sentence from the first review, the second from the second review, the third sentence of the third review, etc. In this way we hope to collect the full spectrum of possible sentences, opinions and aspects which can be found in attraction reviews. Of course, you can skip sentences which are not processed correctly, sentences in other languages, etc. The annotated sentences also include non-opinionated sentences which can be marked at tag level 1 with the label NON-OPINIONATED.

You are asked to annotate **3,000** sentences per language and to keep track of the number of annotated sentences yourself.

5.2 Aspects and examples (Tag 3)

ASPECTS	EXAMPLES
3;Accessibility (car/ public transport)	Highway
3;Atmosphere	Like a Parisian brasserie
3;Animal Well-being	Cages, animal shelters
3;Bathroom	
3;Beds	
3;Beverage	Wine list; French wines
3;Breakfast	
3;Car parking	
3;Cleanliness	
3;Collection	Pictures, sculptures , animals (in a zoo), books, play equipment, etc.
3;Crowds and Queues	(je kunt zo) doorlopen; wachttijd
3;Facilities	
3;Food	Appetizer, desert, menu choice, cheese, gourmand
3;For children	Small children, young people
3;Interior/exterior	Tables, chairs, entrance
3;Internet	
3;Location	
3;Noisiness	
3;Other guests	
3;Reservation/check-in/check-out	

3;Restaurant	
3;Rooms	
3;Safety and Security	
3;Staff	
3;Swimming pool	
3;Value-for-money	
3;Value of Nature	Surroundings
3;Weather	Raining, warm, sun shine

5.3 Type of attraction (Tag 5)

This tag applies to the whole review. It can be given at the first position (w1) . The tag refers to the type of attraction this review is about. The tag ‘unknown’ can be used if the type of attraction type cannot be inferred from the review text (e.g. *it was great, we certainly go back*).

5;amusement park

5;beach

5;bar/dancing

5;church/castle

5;market/shopping

5;memorial site

5;museum

5;theater

5;restaurant

5;zoo/aquarium

5;tour (city/landscape)

5;unknown

Attention with respect to the single-sentence-approach: the type of attraction must be given after reading the whole review (not only the annotated sentence).

5.4 Annotating non-opinionated sentences

If a sentence is not opinionated , the whole sentence should be labeled at tag level 2 with the table NON-OPINIONATED.. All other tags can be left empty.