ACTER Terminology Annotation Guidelines

The following terminology annotation guidelines have been created within the framework of Ayla Rigouts Terryn's PhD project on automatic terminology extraction (D-TERMINE: Data-driven Term Extraction Methodologies Investigated). She is a member of the LT3 Language and Translation Technology Team at Ghent University, and her PhD is funded by a scholarship from the Research Foundation Flanders (FWO).

The annotation guidelines have been continually updated since 2016, with the current version dating from March 2021. The ACTER dataset (Annotated Corpora for Term Extraction Research) has been annotated according to these guidelines and is used for the first edition of the TermEval shared task. Several publications have appeared which mention the dataset and the guidelines, most notably (Rigouts Terryn, Drouin, et al., 2020; Rigouts Terryn et al., 2018; Rigouts Terryn, Hoste, et al., 2020).

While these guidelines have been created to allow transparent and detailed term annotation with high inter-annotator agreement, they are of course but one of many possible interpretations of the task. Still, we have found that these guidelines work relatively well and, by making them publicly available, hope to be both transparent about our own annotations of the ACTER dataset, and to inspire other researchers for new term annotations. The guidelines are largely language- and domain independent, so can easily be re-used.

The annotations were performed with the BRAT rapid annotation tool (Stenetorp et al., 2011) and the examples provided in this document are often screenshots of annotations made with that tool.

A final note worth mentioning on the guidelines, is that they were created with automatic term extraction in mind. This technology is meant to assist terminologists in keeping up with quickly evolving terms in increasingly more specialised domains. Since automatic term extraction does not aim to extract only the well-established terms (which already appear in term bases), but also new (perhaps still infrequent) terms or non-standard variants of terms, the annotation guidelines take a very descriptive, rather than prescriptive approach. If a linguistic unit (one or multiple words) is used in the text as a term, it is annotated, regardless of whether it is the best term to express that concept.

1 Corpora

These guidelines were originally developed to annotate the ACTER corpora but can be used for annotation of other specialised corpora as well. The ACTER corpora consist of several corpora in 4 specialised domains (corruption, dressage, heart failure, and wind energy) and 3 languages (English, French, and Dutch). A dedicated paper describes the dataset in more detail (Rigouts Terryn et al., 2019) and the latest information can also be found on the TermEval 2020 website (https://termeval.ugent.be) or in the readme.md file of the dataset (https://github.com/AylaRT/ACTER). The current guidelines can be used to annotate any specialised corpus with a clearly defined domain.

2 Annotation Scheme

2.1 What are terms?

There are many different interpretations of what constitutes a term. The Oxford dictionary provides the following definition: "A word or phrase used to describe a thing or to express a concept, especially in a particular kind of language or branch of study." A common example in the medical domain could be "influenza". Laypeople would simply say they have the "flu", whereas specialists would refer to this concept as "influenza".
Other definitions in scientific literature include:

> "The information in scientific and technical texts is encoded in terms or specialized knowledge units, which are access points to more complex knowledge structures. Underlying the information in the text are entire conceptual domains, which are both implicitly and explicitly present, and which represent the specialized knowledge encoded." (Faber & López Rodríguez, 2012, p. 9)
>
> "... terms constitute a subcomponent of the lexicon of a language, since a speaker's competence cannot exclude a specialized vocabulary … terminology is an interdisciplinary field of enquiry whose prime object of study are the specialized words occurring in natural language which belong to specific domains of usage." (Cabré, 1999, p. 32)
>
> "Definitions of "term" often focus on the link between a linguistic unit and a domain concept, …" (Bernier-Colborne & Drouin, 2014, p. 54)

However, since definitions like these are not necessarily practically helpful in deciding whether a given linguistic unit in a text is a term, the following annotation scheme and guidelines have been developed to aid annotators.

2.2 Two criteria, three term types

We define termhood based on two criteria:

1. Lexicon-specificity
2. Domain-specificity

Lexicon-specificity indicates whether a lexical unit is part of common language or if it is only known by specialists. Common vocabulary is not lexicon-specific, the vocabulary specific to experts is.
Domain-specificity indicates whether a lexical unit is relevant to the researched domain. If it is, it is domain-specific. If is unrelated to the relevant domain, it is not domain-specific.
Combining these indicators in a matrix leads to three term labels, as seen in Figure 1.

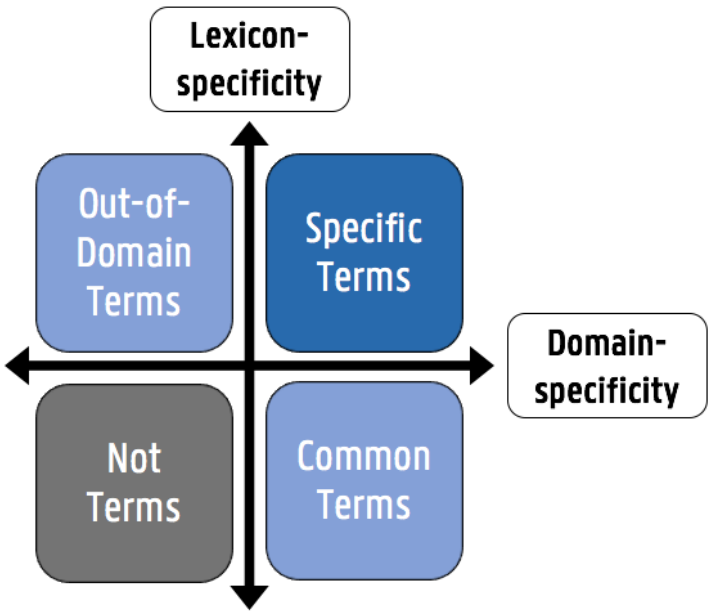

Figure 1: annotation scheme based on two parameters that identify three term labels

Specific Terms are terms according to the strictest definitions of the word: vocabulary only known by specialists in the relevant domain.
Out-of-Domain Terms are lexicon-specific and, therefore, not part of common vocabulary, but they aren't domain-specific, so they aren't relevant to the subject.
Common Terms are the opposite: not lexicon-specific (part of common vocabulary), but domain-specififc (relevant to the subject).
If a lexical unit is neither lexicon-specific nor domain-specific, it isn't a term.

Some examples in the domain of heart failure as placed in their respective categories can be seen in Figure 2.

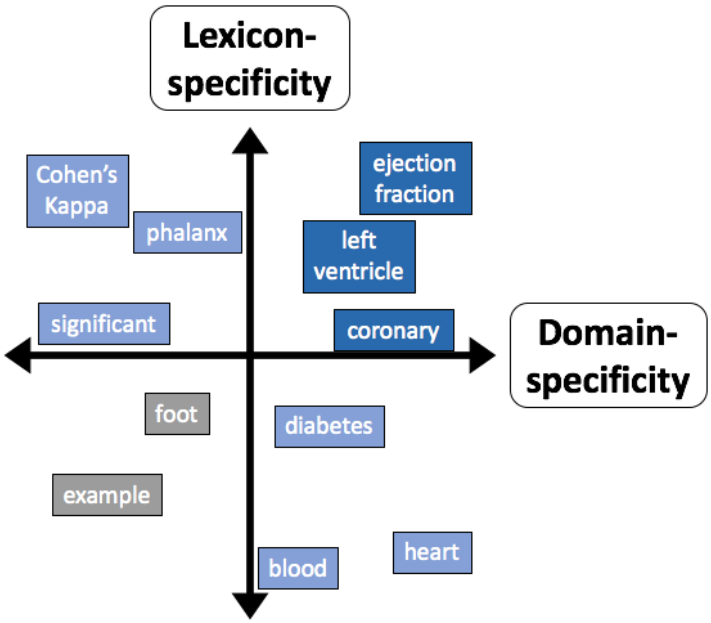

Figure 2: annotation scheme with examples in the domain of heart failure placed along the two axes

Discussion of examples:

- *ejection fraction*: only used by specialists, not understandable for laypeople, strong link with heart failure
- *left ventricle:* used by specialists and strong link to the domain; less lexicon-specific than ejection fraction because more laypeople have at least a vague notion of what a left ventricle is, than of ejection fraction
- *coronary:* lexicon- and domain-specific, even when laypeople may be somewhat familiar with the word, most probably have only a very vague idea of its meaning
- *Cohen's kappa:* statistical term that is not relevant to the domain of heart failure but is not common vocabulary either
- *phalanx*: though this is a lexicon-specific term in the medical domain, the term for a bone in your finger cannot be considered relevant to heart failure in any way, so it is considered an Out-of-Domain Term
- *significant: w*hile *significant* is part of common vocabulary as a synonym of *noteworthy*, it acquires a different, stricter meaning in the field of statistics, namely *probably not due to chance*; because the meaning is different in the specialised domain, it is lexicon-specific
- *diabetes*: understandable for most laypeople and often linked to heart failure
- *heart:* very strong link to the domain of heart failure and common vocabulary; domain-specialists have a much more advanced knowledge of what a heart is, but laypeople are familiar with the same general idea of the heart as an organ that pumps blood
- *blood:* common vocabulary and clear connection to the subject (though in a more general sense)
- *foot:* common vocabulary and, though it could be relevant in some other branch of the medical domain, not relevant for heart failure
- *example:* neither lexicon- nor domain-specific in any sense

Of course, even with these annotation scheme, there is still a considerable degree of subjectivity and people may still place terms in different categories. Nevertheless, the different labels can allow a more intuitive annotation. For example, you can acknowledge that *cardiomyopathy* and *heart* both belong to the domain of *heart failure*, but you are still able to make a distinction between the two by giving them a different term label. As shown in experiments (Rigouts Terryn et al., 2019), using this annotation scheme increases inter-annotator agreement.

More examples in the domain of wind energy:

Specific Terms score high on both axes: they are both lexicon- and domain-specific.
e.g.: The “nacelle” is a part of a wind turbine. It is lexicon-specific because non-specialists generally do not know the meaning of the word; you would not find it in a magazine or newspaper without an explanation. Only people who are familiar with wind turbines will be able to correctly identify the nacelle. It is also domain-specific: the nacelle is part of a wind turbine, which is one of the most commonly used means to generate wind energy, so “nacelle” is undoubtedly relevant to the domain of wind energy.

Out-of-Domain Terms are lexicon-specific, but not domain-specific.
e.g.: The word “orismology” means the identification, specification and description of technical terms. This term is lexicon-specific, because it is not part of the general vocabulary and only specialists would know it, but it is not domain-specific to the domain of wind energy, as it has nothing to do with that subject.

Common Terms are the opposite: not lexicon-specific but domain-specific.
e.g.: The word “wind” is not lexicon-specific: everyone with a basic knowledge of the English language knows the word and its meaning. It is, however, domain-specific, since “wind” is a crucial part of the domain of “wind energy”. The same would be true for words like “sustainable” or “windmill”.

2.3 Named Entities

Apart from the three term labels, **Named Entities** (NEs) were also included, since they can be closely related to terms and ATE is often combined with Named Entity Recognition (NER). On Wikipedia, they are described as “a real world object such as persons, locations, organizations, products, etc., that can be denoted with a proper name” (https://en.wikipedia.org/wiki/Named_entity).

More instructions on the annotation of Named Entities can be found in section 3.3.4

2.4 Four labels

In total, there are 4 labels: 3 types of terms and Named Entities

1. **Specific Terms**
2. **Common Terms**
3. **Out-of-Domain Terms**
4. **Named Entities**

More detailed information on the annotation process, including detailed instructions and tips & tricks for specific difficulties, are provided in the next sections.

3 Annotation guidelines

3.1 General principles

3.1.1 Simple and complex terms

Terms have no minimum or maximum length

Terms can be single words (e.g., *coronary*), two words (e.g., *ejection fraction*) or any number of words (e.g., *terminal deoxynucleotidyl transferase-mediated deoxyuridine triphosphate nick end labelling*). Most terms are no longer than a couple of words, so be careful when annotating very long terms to make sure that it is indeed one comprehensive term and not a combination of different terms.

Terms should be annotated recursively

This means that, if part of a complex term is a term itself, both the longest possible term and the shorter term(s) should be annotated. For example, in the field of automatic term extraction (ATE), *term extraction* and *gold standard* could both be terms, so would have to be annotated as such. Additionally, *term* can be considered a term on its own, so it should be annotated separately. However, *gold* and *standard* are not terms when not combined, so they should not be annotated separately and only when combined.

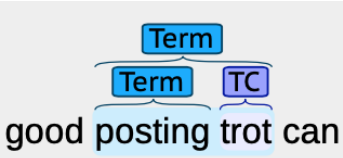

Figure 3: example of recursive annotation single- and multiword terms in BRAT

The example in Figure 3 shows a recursive annotation in the domain of dressage. In this case, *posting trot* has been annotated as a Specific Term, while the parts, *posting* and *trot* have been annotated separately as well. The two single-word terms have specific meanings both separately and combined.

The full term and its parts do not always have the same label

For example, in Figure 3, the full term (*posting trot*) has been annotated as a Specific Term and so has the first part (*posting)*, while *trot* was annotated as a Common Term, since *trot* is considered part of the general vocabulary, while most people do not know what *posting* means in the context of dressage.

Do not annotate below word level

Only annotate a unit if there is a whitespace or a "-" between the words. For example, if you were to annotate *football*, you do not *annotate* foot and *ball* separately, because

they are part of the same word. However, in a term like *angiotensin-converting enzyme*, *angiotensin* may be annotated separately because there is a dash that separates it from the next word.

Pay special attention in the case of single apostrophes. When they are used for a plural (as they often are), the entire plural form should be annotated (e.g., *NSAID's* should be annotated only as a whole). When the apostrophe is part of an English possessive, it should not be annotated (e.g., with *patient's*, only *patient* should be annotated, without the *'s*). In French, the apostrophe is often used in combination with articles or pronouns (e.g., *l', d', s')* and should not be annotated, unless it is a reflexive verb which is a term in its reflexive form. In conclusion, the apostrophe is only part of the term when it is inherently part of the term (plural or reflexive verb), not in other cases.

4.1.2 Split terms

Sometimes complex terms are "split": they are interrupted by other words or punctuation. This can happen for example in ellipses (e.g., in the phrase *A- and B natriuretic peptides*, both *A natriuretic peptides* and *B natriuretic peptides* are present as terms, but the former is interrupted) or when part of the term is between parentheses (e.g., in the phrase *angiotensin converting enzyme (ACE) inhibitors* the full term *angiotensin converting enzyme inhibiters* is interrupted by the abbreviation).

The problem with these split terms is that most annotation tools only allow annotations of consecutive words. To circumvent this problem, we included a "relation" label, which allows the annotator to link two parts of a discontinuous ("split"") term and label the relation. This way, in our example, you could annotate *ACE* and *inhibiters* separately as terms and link them with a "Split Term" relation. We also included "Part of" labels, when one of the parts of a split terms is not a term in its own right. For instance, in the former example of *A- and B natriuretic peptides*, the letter *A* is a "Part of Term", and *natriuretic peptides* is a term itself; the link between the two is "Split Term". When annotating with the BRAT online annotation tool, the result is displayed in figure 4.

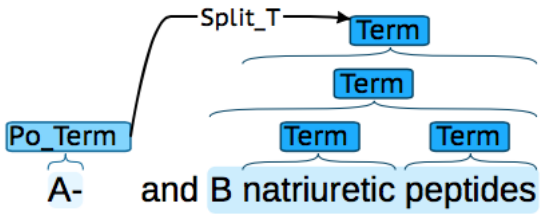

Figure 4: Split Term in BRAT

In this example, the split term is *A natriuretic peptides*. The first part of this term (*A*) is not a term on its own, so it gets the "Part-of-Term" label. In contrast, *natriuretic peptides* is a term on its own, so it does not require the "part of" distinction.
This is also the only case where annotations can be made below token-level, especially in Dutch when there is an ellipsis that includes a compound term (e.g., *hart- en nierfalen* is a combination of *hartfalen* and *nierfalen*, so *falen* can be annotated as a part of a term, even though it is below token-level. However, these should always

be “part-of” annotations and should never be annotated with either of the four main labels.

Since these additional labels are available in all term types and as Named Entities, this means there are 8 labels in total and 4 relations:
Normal labels:

1. **Specific Term**
2. **Common Term**
3. **Out-of-Domain Term**
4. **Named Entity**
5. **Part of Specific Term**
6. **Part of Common Term**
7. **Part of OOD Term**
8. **Part of Named Entity**

Relation labels:

1. **Split Specific Term**
2. **Split Common Term**
3. **Split OOD Term**
4. **Split Named Entity**

3.1.3 Untranslated terms

An attribute “**untranslated term**” (see figure 5) was added to indicate terms in languages different from the rest of the text. For instance, in Dutch, English terms might be used sometimes. This attribute can be added to any term label. You can select “**English**”, “**French**” or “**Other**”. Neo-classical (Greek or Latin) terms are not considered untranslated and should not get this extra attribute. It should be noted that this label is not used for Named Entities (which are even more often in a different language).

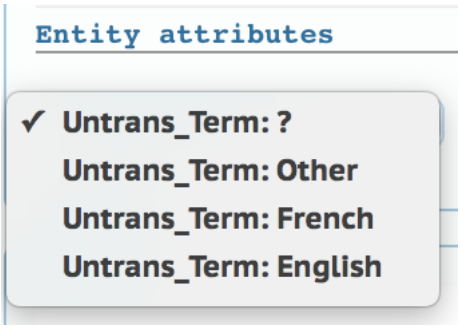

Figure 5: Untranslated terms

3.1.4 Part-of-Speech

All (combinations of) content words can be terms

Even though there is a lot of discussion about the possible parts-of-speech for terms, we decided **all (combinations with) content words are possible terms**: nouns, adjectives, verbs and adverbs. The proportion of nouns will probably be much larger than the others, but there is no valid reason to exclude them.

3.1.5 Consistency

Each occurrence of every term and Named Entity should be annotated and assigned the same label (unless they are used in different ways).

The corpus will contain many recurring terms and Named Entities. Every instance of every annotation should be annotated. If a text about heart failure contains the word *cardiomyopathy* 42 times, it should be annotated with the same label 42 times. If it appears both in singular and plural form, with and without capitalisation, it should still be annotated every time. Try to be **logical**: if the noun form of a word is a term, then the adjective form should probably be as well. For example, if you annotate the noun *aerodynamics* as a Common Term, then the adjectives *aerodynamic* and *aerodynamical* should be annotated as Common Terms as well.

Exceptions are of course possible for words/terms with multiple meanings: if the same term can have multiple meanings, it may be that it has different labels for those meanings. For example, in the dressage corpus, *collection* is a Specific Term, so all variations of that term are annotated as such, e.g., *collect, collected, collects*. However, when it occurs in a non-terminological meaning or in a different corpus where it is not a term, it should not be annotated.

Consistency can also count **across languages**. If you annotate a term in one language, there is a good chance you should annotate the equivalent in other languages with the same label. There are exceptions to this rule, e.g., the English term *pneumonia* could be considered a Common Term, since most non-specialist people are familiar with it. However, in Dutch, *pneumonie* might be a Specific Term, because non-specialists would use *longontsteking* (literally: lung inflammation) and the neo-classical *pneumonie* is a lot more specialised.

3.2 Practical tips

3.2.1 Specific Term or Common Term?

One way to distinguish between Specific Terms and Common Terms is to ask yourself if the word or phrase would occur (without explanation) in **popular media** (magazines, newspapers, …) that are aimed at a large, general audience. Terms used in that context (without any explanation) are intended to be understood by most of the non-specialised readers. So, if the term appears in popular media, it would be reasonable to assume it is a Common Term. You can use Google News to check how the term is used in news sources, but make sure the sources are aimed at a general audience.

An example where it is difficult to decide between Specific or Common Term could be the term *heart failure*. To check, type in “heart failure” (between quotation marks” in Google News Search (see figure 6). Do not simply trust the number of hits, because some of the hits may be from specialised journals, see for example figure 7.

Figure 6: using Google News to decide between the Specific and Common Term labels

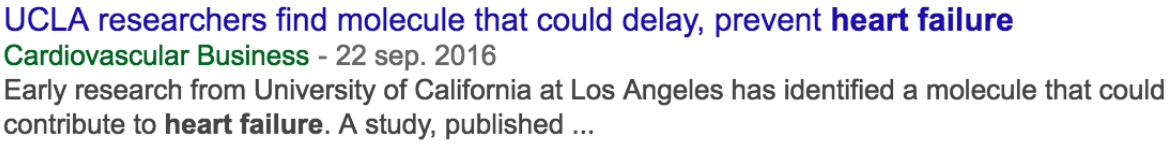

Figure 7: Google News results

Look for popular media sources that aim at large audiences, for example Fox news. You can even add this to the search, for example: “heart failure” “Fox news” to force Google to look for hits in that source. In our case, we get 2.100 hits and a closer look reveals that this term usually doesn’t get an additional explanation. Therefore, we decided to give it the label “Common Term”. Of course, this is only one strategy and it is not fool proof, but it can help you to decide when in doubt.

3.2.2 Specific Term or Out-of-Domain Term?

To recognise Out-of-Domain terms, always keep in mind the how the domain of the corpus is defined. If you think the term would be likely to **occur more often in texts about the relevant subject than in texts on other subjects**, that is an indication the term is domain-specific. Again, you can use Google as an indication. If you look up the term in combination with the domain name, does this give many matches? However, be very careful with this method, because many words can have several meanings and the number of hits Google gives you are not always very accurate, so use your common sense.

Also remember to distinguish between the larger context of the corpus and the more specific subject. For example, think of the domain of wind energy. Wind energy is a subdomain within the field of energy or technique, and some terms about energy are relevant enough for wind energy that they can be considered terms (e.g., *kinetic energy, electrons*). Others are not (e.g., *yellowcake* (Specific Term in domain of nuclear power)). Another example for the subdomain of heart failure would be terms within the medical domain that are not related to heart failure in any way, such as *osteogenesis imperfecta*. This is a specific medical term, but not at all related to heart failure, so it should be given the label “Out-of-Domain Term”. On the other hand, terms which are used in the entire medical domain, including heart failure, may also be relevant. For instance, *comorbidities* is used in the medical term in general and can be relevant for heart failure as well, though it isn’t specifically related to it. In this case, *comorbidities* can be a Specific Term.

3.2.3 Term at all?

If you are not sure whether a lexical entity should be considered a term at all (not even common or Out-of-Domain term), the best strategy is not to annotate it. If you are hesitating, it probably means the term scores quite low on both specialisation and domain-specificity, so it should not be annotated. In our experience, the more you annotate, the more you think about annotating: "if I consider this a term, then that should probably also be a term". So, the recommended strategy is: **when in doubt, do not annotate**. One additional strategy is to consult Wikipedia: if the entity has its own **Wikipedia** page, chances are high that it is a clearly defined concept, so is more likely to be a term.

Always keep in mind **consistency** as well. For instance, in the corpus on dressage, it was originally decided to annotate *leg* and *hand* as common terms, since they are constantly mentioned in dressage instructions. However, if these body parts were annotated, it was only logical to also annotate other (relevant) body parts for the sake of consistency, even if they do not appear quite as frequently as the former two (e.g., *fingers, back muscles, knee,* etc.).

3.2.4 Search function

For consistency, it can be very useful to use the search function (**ctrl + f**). However, be careful. When using BRAT, depending on which browser you use, ctrl + f will either bring up the browser's search function or BRAT's search function and they have a different functionality. A browser's search function is usually not **case-sensitive**, but BRAT's is. So, if you look for "*example*" with BRAT's search function, it will not find "*Example*". In the standard setting, BRAT's search function is also limited to the full string. This means that, when you type in "*legal*", you will only find "*legal*" and not "*illegal*". You can go to the "advanced options" in the search window and check the "**any substring**" box to change this setting. This can make it a lot easier. For instance, if you want to annotate "*technology*" and type in "techn", you will immediately find all kinds of variations, like "*technologies*", "*technological*", "*technical*", etc. Some words can be labelled differently depending on the **context** (see 4.1.5), so do not blindly annotate every occurrence of a term without checking.

3.2.5 List of difficult cases

It can be helpful to keep a list of difficult cases per language and subject. Once you have decided on a certain label for a term, it is important to annotate this term consistently throughout the corpus. If you hesitated about a label, you may have forgotten what you decided by the time you see the term again. To avoid having to look up the first occurrence again or labelling the same term differently, **write down the terms for which you had trouble deciding** in a separate list which you can easily consult and elaborate while annotating.

Another tip would be to consider consistency across languages in this list as well. If you decide on one label in one language, you can immediately check the translation in a different language to see if the same reasoning applies there. This way, you avoid having to go through the same process for every language and you can ensure a better consistency.

3.2.6 Google and Wikipedia

If you use BRAT, you can immediately click through to search the lexical entity you want to label on Google or on Wikipedia. BRAT will automatically direct you to the English Wikipedia, so make sure you go to the appropriate language. As explained before, googling an annotation can be very helpful and even simply reading the little snippets of the first search results can give you unexpected information about which label to use. Wikipedia is also a great help. You can go to the Wikipedia page of the corpus subject as a sort of domain summary and look at the terms in that document.

3.3 Additional guidelines

3.3.1 Abbreviations

Abbreviations should be given the same label as the full forms

This is not always obvious since abbreviations can be less known (more lexicon-specific) than the full form. However, trying to decide whether this is the case can be very difficult, especially with unofficial abbreviations. Therefore, to ensure consistency, always give the abbreviation the same label as the full form. This counts for both terms and Named Entities.

A difficult case can be when Named Entities are referred to by only a short part of the entire term, but still written with capitalisation. For instance, suppose the *Declaration of Human Rights* appears in the text and is annotated as a Named Entity, but is later referred to as "*the Declaration*". Should *Declaration* be annotated? In such cases, it was decided not to annotate the abbreviated form, since such "abbreviations" are too vague and not even always capitalised. However, there are three exceptions to this: *Parliament*, *Council*, and *Commission*, as referring to the EU institutions. These abbreviated forms for are so commonly used in the context of the corruption corpus, that it was decided the link between these abbreviated forms and the full forms is clear enough to still annotate them as Named Entities.

3.3.2 Website addresses

Some texts may contain website addresses. Even though these links may contain some terms embedded in the address, they **should not be annotated**. Spacing and capitalisation is rarely as it would be in normal text, so annotating these would give a distorted view of the terms.

3.3.3 General nouns

Sometimes adjectival terms are combined with very general nouns, such as: **aspects, things, cases, elements, process, etc**. These nouns are generally not terms and **should not be annotated**. However, the accompanying adjective may still be important, so can be annotated separately. An example in French is *aspects épidémiologiques*, where the adjective can be considered a term but the combination with the noun not.

3.3.4 Named Entities

Do not annotate Named Entities in a different language recursively

Named Entities generally follow all the same rules as the other terms, except that they should not be annotated recursively if they are in a different language. If the Named Entity is in the same language as the rest of the corpus, recursive annotation is no problem, but if it is in a different language, only the full Named Entity (longest possible form) is annotated.

Again, there are no hard-and-fast rules to make the distinction between terms and Named Entities, but an internet search can be helpful. For instance, **Named Entities are not generally preceded by indefinite articles** (a/an) because they refer to a single entity. You would not usually say "an America". You would, however, say "*a Doppler echocardiography*", which can be an indication that, despite the capital letter, *Doppler echocardiography* should be considered a term rather than a Named Entity.

Another important remark in the medical domain is that the names of substances or medicines should not be annotated as Named Entities, except when they are brand names. For instance, *Trastuzumab* is a type of medicine, but as can be read on Wikipedia, it is "sold under the brand name Herceptin". So, while *trastuzumab* may sometimes look like a Named Entity and is sometimes capitalised, only the brand name *Herceptin* is actually a proper name and *trastuzumab* is a Specific Term.

3.3.6 Modifiers (adjectives)

Be careful about including adjectives in an annotation. **Does the adjective belong with the term or is it simply a description?** One way of testing this, is to see whether the adjective have a specific meaning within the domain or not and by investigating how (often) the adjective and noun are used in this combination.

For instance, in the medical domain: should you annotate *mild anemia*? At first glance, the answer would be "no", you only annotate *anemia*, because *mild* describes the *anemia*, but you could just as easily use it to describe the weather, it has no specific meaning in this context. However, when you google *mild anemia*, you find that it is defined as anemia with hemoglobin levels between 10 g/dL and 12 g/dL, therefore, *mild* has a very specific meaning in combination with *anemia* and *mild anemia* can be annotated as a whole (though *mild* in itself is still not a term).

Another example in the medical domain: do you annotate *acute*, which appears in combinations such as *acute pancreatitis*? Wikipedia is helpful in this case, because when you look for *acute* there, you get the page depicted in figure 9. The fact that the adjective has its own Wikipedia page, even with the qualifier *medicine* added, is enough to decide that it has a specific enough meaning within the field of medicine to be considered a Term, or, in this case, a Common Term. Since *acute pancreatitis* is a well-defined form of pancreatitis (e.g., it also has its own Wikipedia page), the resulting annotating could be as shown in figure 8.

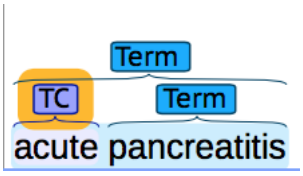

Figure 8: example annotation of multi-word term with terminological modifier

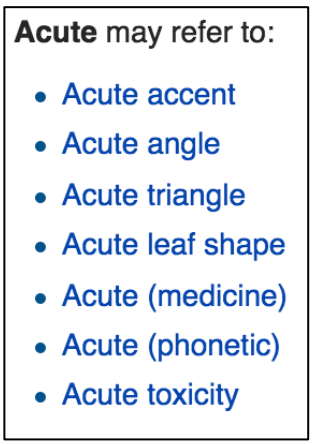

Figure 9: "acute" on Wikipedia

One final example is *severe reductions in LVEF*, again in the domain of heart failure. We can apply a similar reasoning as with *mild*: the adjective does not have a specific meaning in the medical domain, which would be an indication not to include it in the annotation. When googling *severe reductions in LVEF*, one of the first hits reads: "moderate to severe reductions in LVEF (in this study an. LVEF 40%)", indicating that what constitutes *severe* is not absolutely defined, but determined on a study-to-study basis. This time, the conclusion would be not to annotate *severe*.

3.3.7 Units of measurement

Units of measurements (g, m, l, W, P, etc.) must go through the same process as other terms, looking at domain-specificity and specialisation. In general, the same rule can be followed as for abbreviations: if you annotate the full form, also annotate the abbreviation. If you see a unit of measurement such as *kg*, consider whether you would also annotate what it stands for, not only *kilogram*, but also *weight*. Also consider the following two remarks.

Do not include "/". For example, it some contexts, it may be useful to annotate "km" (kilometre) and "h" (hour), but do not annotate *km/h* together. The reason for this, is that these kinds of measurements can get very complicated, so it makes more sense to only annotate the parts that are relevant.

Some of the texts in the corpus are converted .pdf files. The problem with that, is that (mathematical) formulas do not always convert well and what is readable in a pdf may become a jumble of wrong characters in a txt-file. Consequently, you should be very careful about annotating these symbols and follow the previously mentioned strategy: when in doubt, do not annotate.

3.3.8 Spelling mistakes and typos

It is always possible that a term in the corpus is misspelt. The rule here is that you **annotate the misspelt term, as long as it is still easily recognisable**. For example, if you come across *haert failure* instead of *heart failure*, you can annotate this as you would the correct form.

4 Conclusion

The current document can be used as a guideline for term annotation. It is as detailed as possible and continually updated but does not claim to be exhaustive. Even with detailed guidelines, term annotation remains a very ambiguous task and there is not always one “correct” decision. Sometimes, an argument can be made for different annotations, even following the same guidelines. If you are unsure about a certain annotation, try using all means at your disposal; Google and Wikipedia can be very helpful. If you have a recurring problem, please let me know by sending an e-mail to Ayla.RigoutsTerryn@Ugent.be and I will try adding logical and helpful guidelines for that problem to this document.

5 Bibliography

Bernier-Colborne, G., & Drouin, P. (2014). Creating a test corpus for term extractors through term annotation. *Terminology*, *20*(1), 50–73.

Cabré, M. T. (1999). *Terminology. Theory, Methods and Applications* (J. C. Sager, Ed.). Amsterdam/Philadelphia: John Benjamins.

Faber, P., & López Rodríguez, C. I. (2012). Terminology and Specialized Language. In P. Faber (Ed.), *A Cognitive Linguistics View of Terminology and Specialized Language* (pp. 9–32). Berlin: Walter de Gruyter GmbH & Co.

Rigouts Terryn, A., Hoste, V., & Lefever, E. (2018). A Gold Standard for Multilingual Automatic Term Extraction from Comparable Corpora: Term Structure and Translation Equivalents. *Proceedings of LREC 2018*. Presented at the Miyazaki, Japan. Miyazaki, Japan: ELRA.

Rigouts Terryn, A., Hoste, V., & Lefever, E. (2019). In No Uncertain Terms: A Dataset for Monolingual and Multilingual Automatic Term Extraction from Comparable Corpora. *Language Resources and Evaluation*, 1–34. https://doi.org/10.1007/s10579-019-09453-9

Stenetorp, P., Topić, G., Pyysalo, S., Ohta, T., Kim, J.-D., & Tsujii, J. (2011). BioNLP Shared Task 2011: Supporting Resources. *Proceedings of BioNLP Shared Task 2011 Workshop*.