Annotation guidelines for the project *Probing patterns of argumentative discourse*

Lea Kawaletz, Heidrun Dorgeloh, Zeljko Bekcic, Stefan Conrad

Last updated July 22, 2022

Contents

1 Introduction

The overall aim of the project is a better understanding of the relevant linguistic features both for machine learning and for humans when processing argumentative discourse. To this end, we are compiling and annotating a corpus of news opinions about Covid-19-related topics (the *ArguNEWS corpus, Conrad et al. in prep*), which provides us with an inventory of arguments (henceforth *the database*) that we will then use and manipulate as training data. We are thus investigating *arguments* and not *argumentation*, zooming into individual text passages with an argumentative function rather than investigating the overall structure of and relations within an argumentative text.

     The annotation guidelines provided in this document serve to instruct new annotators, and they are currently still being developed as corpus creation proceeds.

2 Procedure

The general procedure is as follows. First, new annotators receive the guidelines and resolve any open questions with the other project members, either along the way or in designated project meetings. Then, they annotate two test texts from the corpus in order to get familiar with the annotation tool and to identify any further questions. Once these are solved, the test annotations are corrected and further texts are unlocked in the annotation tool. Upon launching the tool, annotators are automatically assigned one of the currently unlocked texts. Any questions that arise while annotating are collected and discussed, and the annotation guidelines are adapted if necessary. Once a text has been annotated by three annotators, one of the curators compares their annotations and creates a gold standard.

3 Definitions

Argumentative texts are characterized by presenting a central, disputed issue, the **major claim**, which the author argues for or against (Stab & Gurevych 2017). More precisely, they intend to persuade an addressee to believe or evaluate or do something by providing a number of **arguments** (Eemeren & Grootendorst 2004; Stede & Schneider 2019). Each argument is categorized into one of the basic categories given in Table 1 (see Eggs 2008; Stede & Schneider 2019).

Table 1: Argument categories (based on Eggs 2008; Stede & Schneider 2019)

	epistemic	ethical	deontic
positive claim	x is true	x is good	do x
negative claim	x is false	x is bad	don't do x

An argument consists of two components, the *claim* and the *premise*. The **claim** is a controversial statement which provides the topic of the argument, and its **premise** is a statement which provides evidence or expresses reasoning that either supports or attacks the claim (Stab et al. 2018). The link between a claim and its premise can thus be conceptualized as a directed argumentative relation, with a premise as the source and a (major) claim as its target (Stab & Gurevych 2014b). Any given text passage may have double functions: If a claim attacks or supports another claim or the major claim, it doubles as a premise, and if a premise is attacked or supported by another premise, it doubles as a claim. Such corpus examples enter the database multiple times.

The cover term for *claim* and *premise* is **argumentation discourse unit** (*ADU*). An ADU "plays a single role for the argument being analyzed, and is demarcated by neighboring text spans that play a different role, or none at all" (Stede & Schneider 2019). A given ADU may be shorter than a sentence, provided that it constitutes a proposition of its own, or span multiple sentences at once. On the form side, the three functional categories (*claim, premise, ADU*) are thus instantiated by **text passages** of any length. Note that we do not annotate arguments with unexpressed ADUs, that is, claims or premises which are not stated in the text but which need to be inferred by the addressee (unlike e.g. Stede & Schneider 2019, 15). Relations between them, however, need not be explicitly stated in the text, but may remain implicit.

4 Corpus examples

The examples in this section have been extracted from the ArguNEWS corpus. ADUs are marked as follows:

- major claim

- **claim**

- <u>premise</u>

For improved readability, the argument components are depicted as sub-examples. All claims and premises are adjacent to one another if not otherwise indicated by the inclusion of unmarked text or by '[...].'

The following sections provide examples for the major claim (section 4.1), the three claim categories (section 4.2), different constellations in which claims and premises appear (section 4.4), and text passages that we do not consider in our annotation of arguments (section 4.5).

4.1 Major claim

In the example below, the major claim, 'many people are resisting vaccination,' is expressed in the article's headline in (1a). The text passages in (1b) to (1d) present three claims from the article's body, all supporting the major claim.

(1)   a. `Why So Many People Are Resisting Vaccination`
       b. [...] **Americans born after the mid-20th century belong to the vaccine-spoiled generations.**
       c. [...] **[W]e came to forget our vulnerability and even disregard or grow suspicious of the vaccines that have saved countless lives.**
       d. [...] **[Medical] protection has made us both complacent and risk-averse.**

4.2 Argument categories

4.2.1 Epistemic argument

An epistemic argument contains a claim which states that something is or is not the case. A supporting premise provides evidence or reasoning for the claim, as in (2), while an attacking premise disputes or challenges it, as in (3) (see also Eggs 2008). Thus, in (2), the author argues for two upcoming Covid vaccines, and he supports this by drawing a parallel to an existing vaccine. In (3), the author challenges his claim that vaccination is going well by conceding that the numbers are currently declining.

4

(2) a. **It's possible that vaccines under development by Novavax and Sanofi, which are likely to begin late-phase clinical trials later this year, may be better for the elderly, Dr. Offit noted.**

    b. <u>Those vaccines contain immune-stimulating particles like the ones contained in the Shingrix vaccine, which is highly effective in protecting older people against shingles disease.</u>

(3) a. **Vaccination is going relatively well in this country,**

    b. <u>although the number of people who receive a dose each day is down from its peak.</u>

These examples also show how the relation between claim and premise can either be explicitly stated (*although* indicating an attack relation in (3)), or left implicit, as in (2). The fact that we are here dealing with a support relation needs to be inferred by the annotator, and must be tested by applying the tests described in section 5.4.1.

Table 2 presents a list of linguistic features which may indicate epistemic stance, gathered from our corpus and the literature (Gablasova et al. 2017, refs). Note, however, that these are merely possible indicators and neither necessary nor sufficient features, and also that epistemicity is often not overtly marked in written discourse in the first place (ref).

Table 2: Examples for linguistic features indicating epistemic arguments (adapted from Gablasova et al. 2017)

Category	Examples[1]
Adverbial expressions	actually; apparently; certainly; definitely; evidently; for sure; kind of; maybe; no doubt; obviously; perhaps; possibly; predictably; probably; roughly; sort of; surely; undoubtedly; without a doubt
Adjectival expressions	doubtful; possible; probable; likely; I am certain; I am confident; I am convinced; I am sure; I am certain; I cannot be sure; I cannot be certain;
Verbal expressions	appear; seem; I assume; I believe; I bet; I doubt; I gather; I guess; I mean; I know; I presume; I reckon; I suppose; I suspect; I think

---

[1]The features in this column are representative of all their variants, including contraction (e.g. *I am certain* vs. *I'm certain*), verbal inflection (e.g. *appear* vs. *appears*), modification (e.g. *without a doubt* vs. *without any doubt*), and negation (e.g. *I suppose* vs. *I don't suppose*; *possible* vs. *impossible*).

4.2.2 Ethical argument

An ethical argument contains a claim which expresses that something is positively or negatively evaluated, i.e. judged as good or bad, excellent or horrible, or anything in between (e.g. *fine, ok, fair, poor, very good*). A supporting premise provides reasoning for this judgment, which is exemplified in (4). In the claim, a scientific development elaborated on earlier in the text (referred to by *this*) is positively evaluated as *good*. The premise gives a reason for this evaluation: The development can give people hope. In (5), the author judges the unwillingness of "teachers [...] to return to school" to be fine, as indicated by "I do not blame." The premise, however, attacks this claim by providing conditions that would reverse that judgment.

(4) a. **Still, this is good news, said Dr. Joshua Sharfstein, a vice dean at the Johns Hopkins Bloomberg School of Public Health and a former F.D.A. deputy commissioner.**

  b. He said: "I hope this makes people realize that we're not stuck in this situation forever.[...]"

(5) a. **I do not blame teachers for being unwilling to return to school in places where administrators and local officials have been in denial about Covid-19 or have been unwilling or unable to do this preparatory work.**

  b. But once schools have put in place appropriate science-based steps, most teachers (those not in high-risk groups) should return to their jobs.

Table 3 presents a list of linguistic features which may indicate ethical arguments, gathered from our corpus and the literature (Martin & White 2007, refs).

Table 3: Examples for linguistic features indicating ethical arguments

Category	Examples
Lexical (good)	significant; awesome; interesting/of interest/interestingly; great; splendid
Lexical (bad)	bad/badly; lacking; pathetic; flawed

4.2.3 Deontic argument

A deontic argument contains a claim which demands that something be done or not done. A supporting premise provides reasoning for this demand, as in (6),where the author gives "keeping everyone safe" as a reason why wearing masks should be

mandated. In (7), the author advises that scientists pay attention to rare side effects of vaccination, while admitting that no serious side effects have been observed so far.

(6)   a. **[M]asking should be mandated and enforced.**

        b. It's not just about your individual risk tolerance, but about keeping everyone safe.

(7)   a. To date, no serious side effects have been revealed, and most tend to occur within six weeks of a vaccination.

        b. **But scientists will have to keep an eye out for rare effects such as immune enhancement, a severe illness brought on by a virus's interaction with immune particles in some vaccinated persons**

Example (7) is somewhat tricky to annotate for two reasons. First, while in most of the arguments in our corpus the premise follows the claim (see also ref), here the order is reversed. Additionally, the connective *but* may lead annotators to conclude that (7b) is an attack on (7a), while, in fact, the opposite is the case. This can be concluded by applying the tests described in section 5.4.1.

Table 4 presents a list of linguistic features which may indicate ethical arguments, gathered from our corpus and the literature (ref).

Table 4: Examples for linguistic features indicating deontic arguments

Category	Examples
Lexical	advisable
Modals of obligation	must; have to; should; ought to

4.3 Support by example

A special kind of empirical evidence that can function as a supporting premise is that of *support by example* (see also Peldszus et al. add ref). A text passage illustrating this is given in (8). Here, the claim is that 2020 was a year with an especially high death rate, which is supported by reporting an exemplary statistic (a given increase for one specific cause of death). That we are dealing with *support by example* here can be straightforwardly tested by applying the pattern 'X. For example, y.' (as in (8')).

(8)   a. 2020 was one of the deadliest years on record for the United States.

        b. Gun homicides and non-suicide-related shootings took approximately 19,300 lives, a 25 percent increase from 2019.

(8′)   a. 2020 was one of the deadliest years on record for the United States.

        b. *For example,* [g]un homicides and non-suicide-related shootings took
            approximately 19,300 lives, a 25 percent increase from 2019.

Such examples are often single data points from a relevant (scientific) study,
or descriptions of similar occurrences from the past or a different location.

4.4 Different constellations

In most of the examples above, we saw a claim followed by a premise. Other
constellations, however, are also more or less commonly found.

4.4.1 n:n relations

One claim is often supported and/or attacked by more than one premise, as in (9),
where the claim has four supporting premises in (9b) to (9e) and one attack in
(9f). So far, we have not encountered examples where one premise supports several
claims in our corpus, which is why we include the constructed example in (10) for
the time being.

(9)   a. **The coronavirus and H.I.V. are different from each other in
           countless ways.**

       b. One is relatively easy to transmit and the other relatively hard.

       c. One lacks the social stigma of the other.

       d. One can kill in weeks while the other tends to kill over time.

       e. Vaccines for one are available for free, while treatments for the other
           can still be prohibitively expensive.

       f. But they both cause deadly infectious diseases that have hit vulnerable
           communities harder.

(10)   a. It is raining outside,

       b. **so we won't be able to go to the zoo today**

       c. **and we should put the covers over the lawn chairs.**

4.4.2 Repeated claim

A frequent argumentative strategy is to state a claim at the beginning of a para-
graph or text, and then to repeat or reformulate the same claim further down. An
example is given in (11), where the claim to leave certain decisions to the experts
is made towards the beginning of a paragraph. Then, the author states who these

experts are, and finally repeats the claim, specifying the involved parties.[2] That is, the named agencies are the experts, while the politicians should take a step back.

(11) a. **First, be ready for anything, and leave it to the experts.**
     b. [...] The major public health and biomedical agencies, such as the National Institutes of Health and its constituent institutes, the Centers for Disease Control and Prevention and the Food and Drug Administration, are led by well-respected scientists and public health professionals.
     b. **The American response should be led by these agencies, rather than politicians.**

In such cases, both claims are annotated as such. Note that several claims may share the same text passage(s) as premise(s), in which case the annotator creates all corresponding relations.

4.4.3 Multiple functions

A text passage may have two functions, serving both as a claim and a premise, as in the constructed example in (12). Here, (12b) functions both as a premise for (12a) and as a claim which is supported by (12c).

(12) a. **Vaccination should be obligatory**
     b. **because it is one of the most effective health interventions available today.**
     c. This has been shown by various studies.

4.4.4 Separated claim and premise

The claim and the premise may not be directly adjacent to one another. In (13), the claim and the premise are separated by further elaborations on the situation. More precisely, the intermittent sentence describes how the information provided in the premise was gathered.[3]

(13) a. **This "the only person you can trust is yourself" mentality has a tendency to cause people to conceive of themselves as individuals and not as citizens.**
     b. Derek Thompson of The Atlantic recently contacted more than a dozen people who were refusing to get a Covid-19 vaccine.

---

[2]For reasons of space, further elaborations have been left out here.

[3]Note that the premise is an argument in itself, so that this text passage enters our data base multiple times in different configurations.

9

c. They often used an argument you've probably heard, too: I'm not especially vulnerable. I may have already gotten the virus. If I get it in the future it won't be that bad. Why should I take a risk on an experimental vaccine?

4.4.5 Syntactically embedded constituents

Insertions may not only happen between a claim and a premise, but also within a single ADU, for example in the form of a relative clause or an apposed noun phrase. In such cases, where the inserted constituent is syntactically embedded, it is annotated as part of the matrix ADU.[4]

The text passages in (14) and (15) (repeated from (7)) are examples for *apposition*, a grammatical construction in which two adjacent constituents (usually noun phrases) are co-referential. In (14), a name is inserted in the middle of the premise, and in (15), *immune enhancement* is explained further as *a severe illness brought on by a virus's interaction with immune particles in some vaccinated persons.*

(14) a. **When coronavirus cases started exploding on the East Coast in March, there were devastating failures by Democratic leaders.**
 b. New York's governor, Andrew Cuomo, forced nursing homes to take back residents who'd been hospitalized for the coronavirus

(15) a. To date, no serious side effects have been revealed, and most tend to occur within six weeks of a vaccination.
 b. **But scientists will have to keep an eye out for rare effects such as immune enhancement, a severe illness brought on by a virus's interaction with immune particles in some vaccinated persons**

4.4.6 Rhetorical questions

While we do not consider arguments with unexpressed ADUs (see section 4.5.3 for examples), we do allow for ADUs which are not explicitly stated, as is the case with rhetorical questions. An example is given in (16), where the claim 'a premature announcement of the Pfizer vaccine could hurt future vaccines' is formulated as an interrogative.

(16) a. **Could a premature announcement hurt future vaccines?**

---

[4]Note, however, that an embedded clause can, in turn, be an ADU, in which case it is in addition annotated as such.

10

b. There's no way at present to know whether the Pfizer vaccine will be the best over all or for specific age groups.

c. <u>But if the F.D.A. approves it quickly, that could make it harder for manufacturers of other vaccines to carry out their studies.</u>

4.5 Negative examples

In this section, we have collected text passages which illustrate what we do not incorporate as arguments or ADUs in our database.

4.5.1 Specification

When identifying premises, annotators also need to sort out text passages which describe the same situation as the claim, merely providing further details. For example, (17a) is a candidate for a claim, but (17b) does not provide evidence or reasoning for it. Rather, it specifies more exactly what happened. This can be tested by applying one of the following patterns (as in (17')):

1. X. What happened is that y.

2. X. More precisely, y.

(17)   a. The U.S. Supreme Court threatens to get into the action, too.

        b. In May, four conservative justices [...] dissented from an order in South Bay United Pentecostal Church v. Newsom allowing California's Covid-19-related restrictions to remain in place for gatherings at places of worship.

(17')   a. The U.S. Supreme Court threatens to get into the action, too.

        b. *What happened is that*, [i]n May, four conservative justices [...] dissented from an order in South Bay United Pentecostal Church v. Newsom allowing California's Covid-19-related restrictions to remain in place for gatherings at places of worship.

4.5.2 Antecedents of anaphora

We do not annotate the antecedent of anaphoric expressions separately. For example, (18a) is the antecedent for *this* in (18b), and we do not mark it as part of the argument. Rather, it is similar to the issue of specification (see section 4.5.1),with (18a) providing further information about what *this* is exactly.

(18)   a. About one in four adults and two out of three children have some fear of needles, and adults may find their fears too shameful to share.

b. **This is a substantial public health problem** <u>because a body of research shows that around one in 10 adults are so afraid of needles that they will delay or avoid vaccinations.</u>

4.5.3 Unexpressed ADUs

We do not include arguments with unexpressed ADUs in our database.

An example for a potential claim which was not included in our database because it is not accompanied by a premise is given in (19). The sentence could be considered to contain two coordinated deontic claims, but there is no premise supporting or attacking them. One possible unexpressed premise, which can be inferred by the reader, is "because these are features that would make sense."

(19)   a. Vaccine passport systems should clarify which shots will be accepted,
       b. and they should be equipped to update immunization requirements when public health guidance changes.

Examples for potential premises without a claim are abundant in the corpus, since any fact could be used to support or attack some controversial statement. An example is given in (20), which could, in theory, support a claim that many people are suspicious of new vaccines, for example.

(20)   a. New vaccines were almost always accompanied by reports of risks and side effects,
       b. and sometimes there were terrible accidents, at least one involving tens of thousands of people made ill by a vaccine that was supposed to protect them.

Such examples would not be beneficial for our database of clear argument examples, and are therefore disregarded.

5 Annotation procedure

This section describes the different steps of the annotation procedure and how the annotations are created in the INCEpTION tool (Klie et al. 2018), available via https://inception.cs.hhu.de.[5] Note that the tool can only be accessed from the HHU network or via the HHU VPN.

---

[5]Further information, including an introductory video and descriptions of the tool's core functionalities, can be found in the INCEpTION User Guide, available at https://inception.cs.hhu.de/doc/user-guide.html#_getting_started.

5.1 General procedure

Upon opening the tool, the annotators are directed either to a random text, or to the text they last annotated. Annotation then proceeds in four steps:

1. Add metadata: The annotator specifies the text with regard to its source, text type, publication date and short title.

2. Identify the major claim: The annotator reads the full text in order to understand the overall argumentation, and annotates or formulates the major claim.

3. Identify claims and premises: The annotator identifies claims and premises according to the criteria delineated in section 5.4.1, and categorizes them.

4. Review and submit: The annotator goes through the whole text again to finalize their annotation, and submits their annotated text.

5.2 Step 1: Add metadata

As a first step, the annotator acquires a very general impression of the text by copying metadata from the document title to the Document Metadata sidebar, which is accessed by clicking the tags icon (box 1 in Figure 1). All texts currently in the corpus share the same source (NY Times, here *NYT*) and text type (opinion, here *op*), so that the annotator only needs to select each from the drop-down menu in box 2, click the '+' sign next to it, and select the respective label from the drop-down menus in box 4. Date and text type need to be entered manually in the corresponding text fields (box 3). Clicking anywhere outside the text field saves the text.
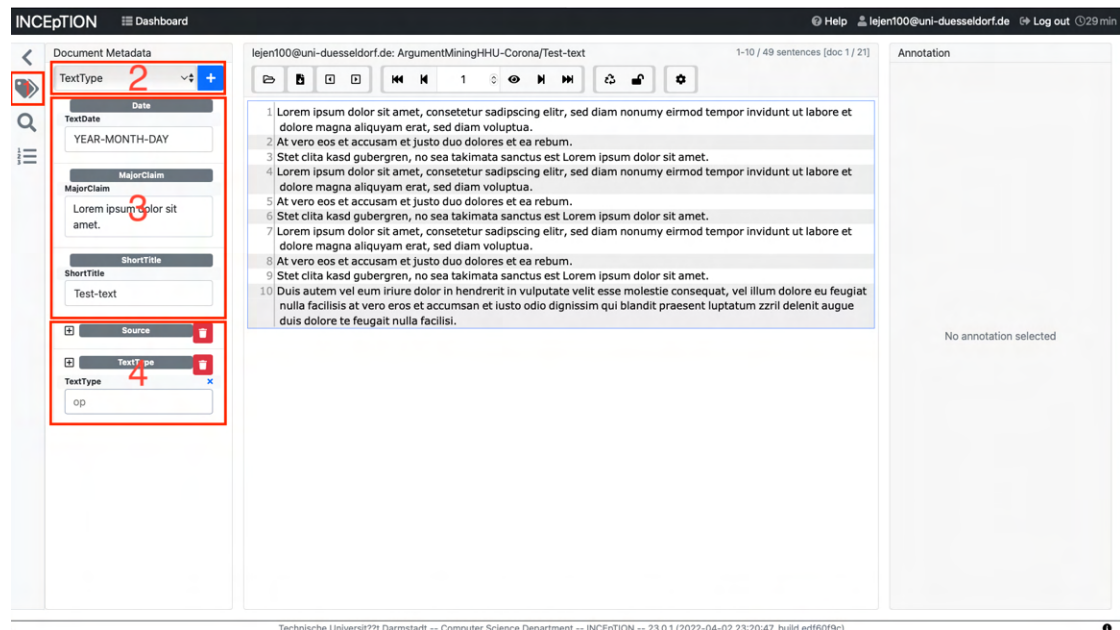
Figure 1: Adding metadata and annotating the major claim

5.3 Step 2: Identify the major claim

Stab & Gurevych (2014a) found that inter-annotator agreement is improved by knowing the topic and stance of a text. Therefore, annotators read the entire text and identify the major claim, which should be an answer to the question *What does the author argue for/against in this text?* (see also section 5.4.1 on testing for controversiality). It may be explicitly stated in the text, for example in the headline of a newspaper article, or inferable from the overall argumentation. If the annotator cannot identify a major claim, the text is not considered further for our corpus.

The major claim is also annotated in the Document Metadata sidebar. Annotators write the major claim, in their own words or as it is formulated in the text, into the corresponding text field (in the middle of box 3 in Figure 1).

5.4 Step 3: Identify claims and premises

In order for a combination of text passages to be included in our database as an argument, it must meet the following criteria (with x and y representing a pair of text passages):

1. x is a controversial statement (the claim)

14

2. x supports, attacks or repeats the major claim

3. x is supported or attacked by y (the premise)

4. x is an epistemic, ethical or deontic claim

This section describes how annotators can ascertain that a given text passage meets these criteria, and how they can create the corresponding annotations in INCEpTION.

5.4.1 Step 3a: Applying paraphrase tests

1. Is x a controversial statement? The status of a given text passage as controversial or not can be tested with the preposed statement *I here argue that*. In (21), this produces a natural statement, while combining it with an uncontroversial fact in (22) is pragmatically odd.

(21) *I here argue that* vaccination should be obligatory.

(22)   * *I here argue that* the sky is blue.

2. Does x support, attack or repeat the major claim? In order to stream-line our annotation process and to create a database with a homogeneous subject matter (i.e.: Covid-19), we only consider arguments which are related to the major claim by a repetition, support or attack relation. If the annotator determines that there is a relation, they annotate the argument as described below, but they do not annotate the argument's relation to the major claim. If there is no relation between the argument and the major claim, the argument is not annotated.

3. Is x supported or attacked by y? If a text passage has been identified as controversial and related to the major claim, the annotator then determines whether it is supported or attacked by another text passage (the premise). This is the case if inserting the two passages in at least one of the following patterns yields (semantically and/or pragmatically) coherent results (where c stands for a potential claim and p stands for a potential premise; see also Table 5 below):

1. For *support* relations:

   (a) c because/since p

   (b) c. This can be seen from the fact that/this is so because p

   (c) c. I claim this to be true because p.

   (d) p, therefore/thus/consequently/which is why/so c

15

2. For *attack* relations:

   (a) c although/even though p

   (b) c, but/admittedly p

   (c) c. However, p

For instance, the two sentences in (23) can be inserted in pattern 1b, yielding the coherent text passage in (23'). Now, the fact that (23b) supports (23a) is made explicit by *this can be seen from the fact that*, and we have determined that (23a) is a claim while (23b) is its premise.

(23) a. That basic sense of peoplehood, of belonging to a common enterprise with a shared destiny, is exactly what's lacking today.

   b. Researchers and reporters who talk to the vaccine-hesitant find that the levels of distrust, suspicion and alienation that have marred politics are now thwarting the vaccination process.

(23') a. **That basic sense of peoplehood, of belonging to a common enterprise with a shared destiny, is exactly what's lacking today.**

   b. *This can be seen from the fact that* <u>[r]esearchers and reporters who talk to the vaccine-hesitant find that the levels of distrust, suspicion and alienation that have marred politics are now thwarting the vaccination process.</u>

Modifications of the two passages that are aimed at preventing ungrammatical or misspelled results (e.g. adapting verb inflection, removing capitalization) are permitted.

It is possible that one of the patterns already exists in the original text passage, as in example (24). In this case, further paraphrases may be tested out for disambiguation, if deemed necessary by the annotator.

(24) a. **Vaccination is going relatively well in this country,**

   b. *although* <u>the number of people who receive a dose each day is down from its peak.</u>

If there are multiple possible premises in the text, the annotator may combine the above patterns with one of the following:

3. For multiple premises with the same polarity:

   (a) Another reason for c is that p2

16

      (b) Another reason against c is that p2

   4. For multiple premises with a different polarity:

      (a) p1. On the other hand, p2.

4. Is x an epistemic, ethical or deontic claim? The final step is to categorize the identified argument into one of the three categories. For this, annotators use the patterns given in Table 5. These patterns make use of the connectors *and* (for support relations) and *but* (for attack relations), of sentential negation (e.g. *not true* negating *true*), lexical negation (e.g. *false* negating *true*), lexical cues (e.g. *approve/disapprove* for ethical claims), and indication of stress by means of italics to increase grammatical acceptability.

Table 5: Patterns testing for claim categories

claim category	positive claim	negative claim
epistemic		
support	and this is true because	and this is false because
	and this is the case because	and this is not the case because
attack	but this is not true because	but this is not false because
	but this is not the case because	but this *is* the case because
ethical		
support	and this is good because	and this is bad because
	and I find this good because	and I find this bad because
	and I approve because	and I disapprove because
	and what is good about this is	and what is bad about this is
attack	but this is bad because	but this is good because
	but what is bad about this is	but what is good about this is
deontic		
support	and do this because	and don't do this because
attack	but don't do this because	but *do* do this because

Linguistic features that point to one of the three claim categories are often also expressed in a given text passage. See Tables 2, 3 and 4 in section 4.2 for examples from the literature and our corpus.

5.4.2   Step 3b: Creating annotations

Once the annotator has identified an argument, they create the corresponding annotation. ADUs are marked by highlighting a text passage, including all punctuation (see box 1 in Figure 2). Note that INCEpTION by default displays a

sentence-oriented editor, which means that each line corresponds to one sentence. This does not imply that each ADU necessarily corresponds to a sentence; they can also be longer or shorter.[6]

Text passages that function as a claim are then annotated with one of the three claim categories by selecting it from the list ADU_Type (box 2). This can be done by clicking the respective label in the drop-down menu, or by using the preset key bindings ($\boxed{1}$ for *DeonticClaim*, $\boxed{2}$ for *EpistemicClaim*, $\boxed{3}$ for *EthicalClaim*, and $\boxed{4}$ for *Premise*). These numbers need to be entered using the number keys on an alphanumeric keyboard; using the numeric keypad will not work.

For text passages that function only as a premise, the annotator selects *premise* from the same list. If a text passage doubles as a claim and as a premise for another claim, the annotator selects the claim type only. As soon as an ADU type is selected, the label above the text passage changes from *ADU* to the selected label.
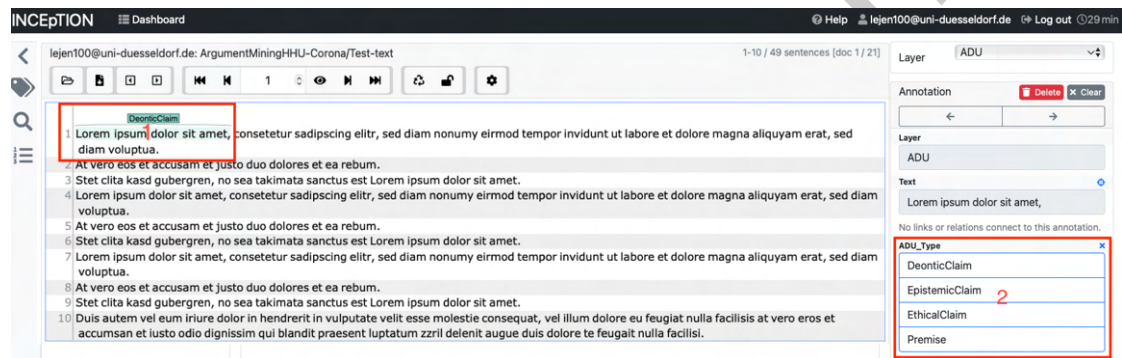

Figure 2: Annotating a claim

Premises are then related to the claim they support or attack by clicking and dragging their label over the claim's label. This creates an arrow from the premise to the claim, labeled *(ADU_Relations)* (box 1 in Figure 3). The annotator then selects *Attacks*, *Supports*, or *SupportsByExample* either by clicking the respective label in the drop-down menu *RelationType* (box 2), or by using the preset key bindings ($\boxed{A}$ for *Attack*, $\boxed{S}$ for *Support*, and $\boxed{E}$ for *SupportsByExample*). Again, the label above the selected text passage changes accordingly.

---

[6]Annotators can change this and other settings by selecting *preferences*, the rightmost button above the editing field.
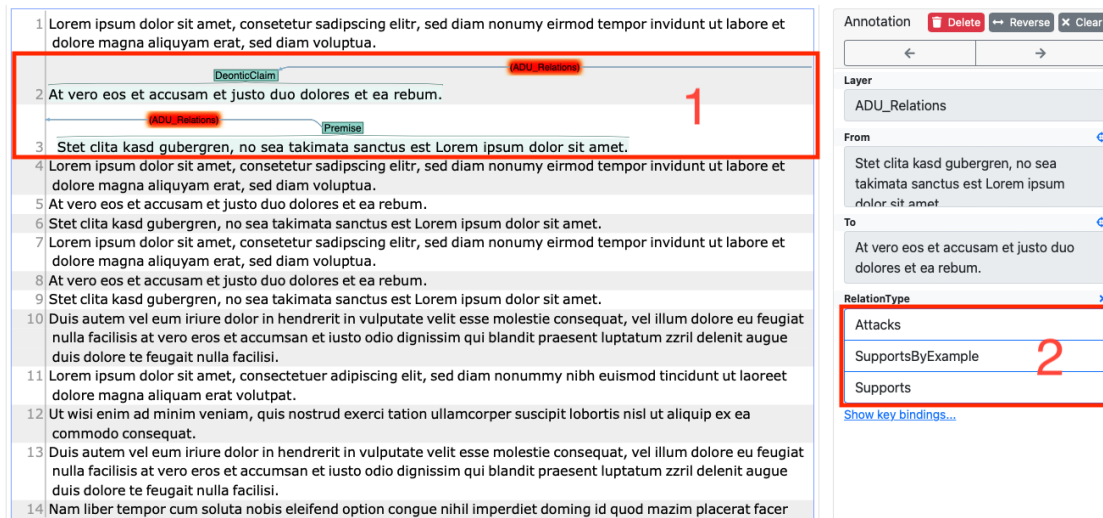
18

Figure 3: Annotating the relation between a premise and a claim

In section 4.4.2, we described the possibility of a claim being repeated. In such a context, it is possible that one and the same premise supports or attacks both instances simultaneously. The annotators then annotate the premise twice, i.e. they mark the text span twice and drag each label onto one instance of the repeated claim (see Figure 4).
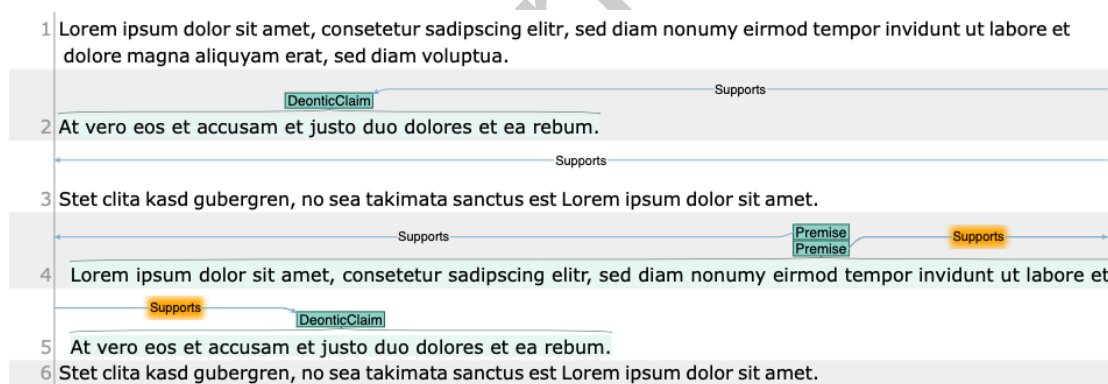

Figure 4: Annotating a premise that relates to a repeated claim

5.5 Step 4: Review and submit

The final step in the annotation process is to go through the created annotations in order to identify mistakes. When they are satisfied with the result, the annotator then submits their file for curation by clicking the *finish document* button (box 1 in Figure 5).
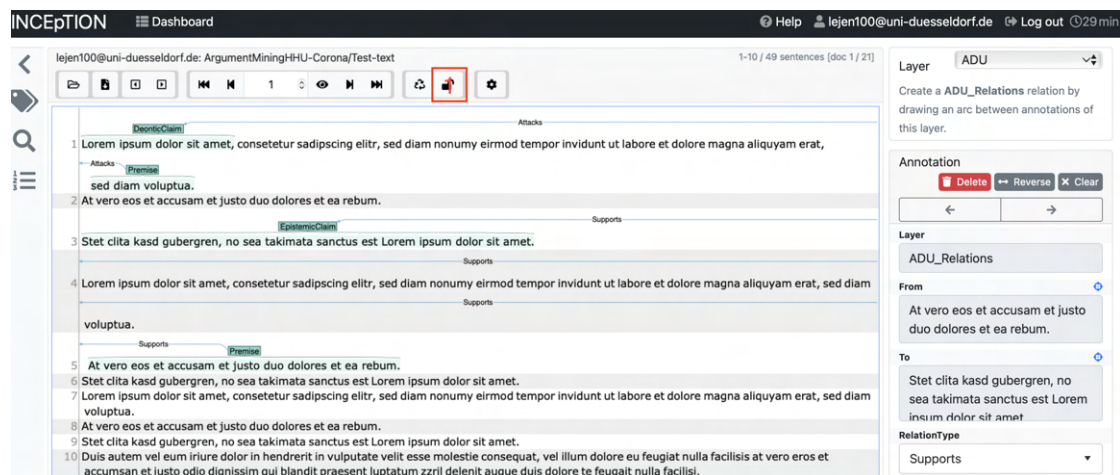
Figure 5: Submitting an annotated text for curation

References

Eemeren, F. H. van & R. Grootendorst. 2004. *A systematic theory of argumentation: the pragma-dialectical approach*. New York: Cambridge University Press.

Eggs, Ekkehard. 2008. 39. Vertextungsmuster Argumentation: Logische Grundlagen. In Klaus Brinker, Gerd Antos, Wolfgang Heinemann & Sven F. Sager (eds.), *Text- und Gesprächslinguistik 1. Halbband*, vol. 16 (Handbücher zur Sprach- und Kommunikationswissenschaft 1), 397–414. Berlin, Boston: De Gruyter Mouton.

Gablasova, Dana, Vaclav Brezina, Tony Mcenery & Elaine Boyd. 2017. Epistemic Stance in Spoken L2 English: The Effect of Task and Speaker Style. *Applied Linguistics* 38(5). 613–637. doi:10.1093/applin/amv055. http://academic.oup.com/applij/article/38/5/613/2952204/Epistemic-Stance-in-Spoken-L2-English-The-Effect.

Klie, Jan-Christoph, Michael Bugert, Beto Boullosa, Richard Eckart de Castilho & Iryna Gurevych. 2018. The inception platform: Machine-assisted and knowledge-oriented interactive annotation. In *Proceedings of the 27th international conference on computational linguistics: System demonstrations*, 5–9. Association for Computational Linguistics. http://tubiblio.ulb.tu-darmstadt.de/106270/. Event Title: The 27th International Conference on Computational Linguistics (COLING 2018).

Martin, J. R & P. R. R White. 2007. *The language of evaluation: appraisal in English*. Basingstoke: Palgrave Macmillan. http://www.myilibrary.com?id=45541. OCLC: 314773660.

Stab, Christian & Iryna Gurevych. 2014a. Annotating Argument Components and Relations in Persuasive Essays. In *Proceedings of the 25th International Conference on Computational Linguistics*, 1501–1510. Dublin.

Stab, Christian & Iryna Gurevych. 2014b. Identifying Argumentative Discourse Structures in Persuasive Essays. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 46–56. Doha, Qatar: Association for Computational Linguistics. doi:10.3115/v1/D14-1006. http://aclweb.org/anthology/D14-1006.

Stab, Christian & Iryna Gurevych. 2017. Parsing Argumentation Structures in Persuasive Essays. *Computational Linguistics* 43(3). 619–659. doi:10.1162/COLI_a_00295. https://direct.mit.edu/coli/article/43/3/619-659/1573.

Stab, Christian, Tristan Miller & Iryna Gurevych. 2018. Cross-topic argument mining from heterogeneous sources using attention-based neural networks. *arXiv preprint arXiv:1802.05758* .

Stede, Manfred & Jodi Schneider. 2019. *Argumentation Mining* (Synthesis Lectures on Human Language Technologies 40). Morgan & Claypool. doi: 10.2200/S00883ED1V01Y201811HLT040.