# Annotating Perspectives on Vaccination

## Roser Morante, Chantal van Son, Isa Maks, Piek Vossen

CLTL Lab, Vrije Universiteit Amsterdam
De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands
{r.morantevallejo, c.m.van.son, isa.maks, piek.vossen}@vu.nl

## Abstract

In this paper we present the Vaccination Corpus, a corpus of texts related to the online vaccination debate that has been annotated with three layers of information about perspectives: attribution, claims and opinions. Additionally, events related to the vaccination debate are also annotated. The corpus contains 294 documents from the Internet which reflect different views on vaccinations. It has been compiled to study the language of online debates, with the final goal of experimenting with methodologies to extract and contrast perspectives within the vaccination debate.

**Keywords:** perspectives, attribution, opinions, claims, vaccination debate.

## 1. Introduction

Debates can be very lively and the language used for debating is often complex and rich. There are, however, not many corpora of real life debates and, if so, the available debates are somehow orchestrated as e.g. parliamentary debates.[1] Nevertheless, the Internet contains a broad variety of texts expressing opinions, beliefs, and statements on many different topics. Partly, we can find these posts on social media platforms such as Facebook or Reddit, but very often they are dispersed all over the web in individual blogs or smaller platforms. The dispersed nature makes it very difficult to grasp the language of online debates and to develop technology that could get to grips with social developments.

In this paper we present the Vaccination Corpus,[2] a corpus of texts related to the online vaccination debate that has been annotated with information about perspectives. The corpus contains documents from the Internet (news, blogs, editorial, governmental reports, science articles), which reflect different views on vaccination. It has been compiled to study the language of online debates, with the final goal of experimenting with methodologies to extract and contrast perspectives within the vaccination debate.

Though vaccines have been proven to be efficient in preventing illnesses, doubts about the reliability of vaccines can easily grow among the population under the influence of negative information about vaccines or unbalanced reports of vaccine risk (Betsch and Sachse, 2012). Since the Internet is a source of health related information for an increasing number of individuals, the spread of information about vaccines and related attitudes can have a high impact on the decisions that the population takes and, specifically, on the decisions of parents as to vaccinate their children or not. The information spread through the internet is influencing the attitude towards vaccination, causing the vaccination rates to drop and outbreaks to appear. For example, an outbreak began at Anaheim, California's Disneyland in

early 2015. At its peak, the outbreak caused 113 children in multiple states to develop measles. The Disneyland measles outbreak and other recent outbreaks of vaccine-preventable diseases in the United States have been associated with substandard vaccination compliance (Majumder et al., 2015; Phadke et al., 2016). In Europe, an increase in measles outbreaks as well as a decrease in vaccination percentages have been shown to be related to behavioral changes in the online interest in terms related to measles and the anti-vaccine movement (Mavragani and Ochoa, 2018).

As the vaccination debate has impact on decisions that affect public health safety, analyzing the way people engage in this discussion and the beliefs that people hold is interesting to understand the situation and act upon it. In order to successfully intervene and curb the spread of preventable diseases due to low vaccination rates, health practitioners need to be adequately informed on public perception of the safety and necessity of vaccines.

The vaccination debate centers around issues related to the safety of vaccinations, the side effects, and the moral aspects of enforcing mandatory vaccination (Wolfe and Sharp, 2002; Mollema et al., 2015). Participants in the debate show strong attitudes in favor and against vaccines. The debate has gained prominence on the Internet due to the increased activity of the anti-vaccination movement, which spreads controversial and uninformed claims about vaccines. The movement became active after the publication in 1998 of a now-retracted paper that linked the MMR vaccine with autism.

One of the characteristics of the debate is that opinions are spread in a myriad of documents on the Internet. Studying how perspectives are expressed in these documents can shed light on how to automatically extract them. For example, it can be useful to know what claims are being made, what opinions and emotions are expressed and who are the sources, i.e. who hold the attitudes.

We understand perspective as a relation between the **source** of a statement (i.e. the author or another entity introduced in the text) and a **target** in that statement (i.e. an entity, event, or (micro-)proposition) that is characterized by means of multiple perspective values expressing the **attitude** of the source towards the target (van Son et al., 2016). All per-

---

[1] https://www.clarin.eu/resource-families/parliamentary-corpora.

[2] The annotations of the Vaccination Corpus are available at https://github.com/cltl/VaccinationCorpus.

spective values together can be seen as a multidimensional characterization of the subjective relation between a source and target. These perspective values may express, for instance, the sentiment (e.g. positive, negative) and the emotion (e.g. happy, sad, angry) towards the target, or the level of commitment towards the factual status of the target but also whether the statement is placed in the future (irrealis) or the past (realis). For example, in Sentence 1 the author attributes to "the establishment media"(source) a positive opinion (attitude) towards vaccines.

1. [The establishment media]$_{source}$ is desperately pushing the myth that [vaccines]$_{target}$ are [completely safe and effective]$_{attitude}$ and have even suggested the government should force you to take them.

We have annotated the Vaccination corpus with three types of information that encode perspectives: attribution, claims, and opinions. We have also annotated vaccination related events in order to be able to extract perspectives on events. We release the annotations so that they can be used to further analyse the phenomena annotated or to develop and test perspective extraction systems, as well as the annotation guidelines.

In Section 2., we introduce the corpus. Section 3. deals with the event annotation, Section 4. with attribution, Section 5. with claims, and Section 6. with opinions. In Section 7. we present some conclusions.

## 2. The Vaccination Corpus

In creating the corpus, an initial set of 50 documents was manually selected to guarantee uniformity in topic but diversity in sources (news articles, activist blog posts, etc.) and manually assessed stance (pro, anti, neutral). The topic was the 2015 measles outbreak linked to the Disneyland Resort in California.Although a source has never been identified, the outbreak likely started from a traveler who became infected overseas with measles, then visited the amusement park while infectious. The outbreak sickened at least 147 people spread across seven states in the U.S. as well as in the neighboring countries Canada and Mexico. It triggered a heated debate where both the pro- and anti-vaccination camps laid the blame on each other for the event. Ceolin et al. (2016b), Ceolin et al. (2016a) and Maddalena et al. (2018) have used this initial set to collect information quality assessments regarding Web documents on the vaccination debate from experts and the crowd.

This initial set was later supplemented with documents that were partly manually selected on the additional topic 'Vaxxed' (a 2016 American pseudoscience documentary film directed by discredited anti-vaccine activist Andrew Wakefield and withdrawn from the 2016 Tribeca Film Festival after widespread criticism), partly automatically collected from the Web using search queries (including terms like *measles*, *disease*, *virus*, *infection*, *chickenpox*), and partly selected from the ControCurator dataset.[3] The data was filtered using a set of keywords (e.g. *vaccin*, *inoculation*) to make sure we only included relevant documents.

The resulting set consists of 294 English web documents (23,467 sentences and 528,727 tokens) from different sources, including Wikipedia, news, editorials, blogs, (pseudo-)science and a variety of health information dissemination websites from both official state-supported institutions (e.g. Centers for Disease Control and Prevention, National Health Service, World Health Organization) and from independent organizations (e.g. Centre for Research on Globalization, National Vaccine Information Center).

| Documents | Sentences | Tokens |
|-----------|-----------|---------|
| 294 | 23,467 | 528,727 |

Table 1: The Vaccination Corpus.

To ensure future accessibility of the web documents, we made use of their archived versions in the Internet Archive (http://archive.org). We took the most recent snapshot in the Archive at the time of collecting, and retrieved the meta data and texts from this snapshot.

All texts in the corpus have been automatically preprocessed with Stanford CoreNLP (Manning et al., 2014)[4] for tokenization, sentence splitting, part-of-speech tagging, lemmatization and dependency parsing. In addition, we used the AllenNLP tools (Gardner et al., 2018)[5] for semantic role labeling, which is a reimplementation of a deep BiLSTM model (He et al., 2017) and uses the PropBank representation (Palmer et al., 2005) of predicate-argument structures in sentences. The preprocessing information has not been used to support the manual annotation process, but we it will be used in our future research to extract perspectives.

Manual annotations were performed with the open source annotation tool eHOST,[6] which also provides options to calculate pairwise inter annotator agreement (IAA) in terms of F-Score, which is the weighted harmonic mean of precision and recall calculated as 2*((precison*recall)/precision+recall). The IAA scores presented in this paper have been calculated with eHost, with lenient span matching in order not to penalize disagreements due to details such as punctuation. In the sections that follow we describe the annotation layers.

## 3. Events

The annotation of events does not aim at identifying all events, but events related to all aspects of the vaccination debate. We adopt the definition of events from the Richer Event Description (RED) Guidelines v1.7 (Styler et al., 2014), where an event is defined as "any occurrence, action, process or event state which deserves a place upon a timeline, and could have any syntactic realization". The annotation guidelines, which are released with the corpus, are based on existing guidelines: TimeML (Pustejovsky et

al., 2006), NewsReader (Tonelli et al., 2014), ECB+ (Cybulska and Vossen, 2014), ISO-TimeML Framework (ISO, 2008), and Richer Event Description (RED). [7] We briefly summarize their content.

All linguistic realizations of (temporally relevant) event mentions should be annotated as events. The annotation of states is restricted to temporally bounded states, thus excluding permanent properties. All verbal predicates, including those that denote a state, give rise to events. Both finite and non-finite verb forms are annotated. Mentions of aspectual verbs are not to be annotated when they contribute to signaling grammatical information concerning the boundedness of an event.

Events realized by nouns can correspond to nominalizations, event nouns, contextual event readings or implicit events. An extension with respect to TimeML and ISO-TimeML is the annotation of events realized by proper nouns (e.g. *World War II*). For this kind of events, the entire noun phrase realizing the event description must be annotated. Event-denoting nouns and present participles in pre-modifier position are never to be annotated as event mentions. Similarly to verbs, aspectual nouns are not annotated as events. Nevertheless, some nouns which may function as an aspectualizer of an event (e.g. "outbreak") must be annotated if subsequently in the text they are (coreferentially) referred to as proper events. Consider this example:[8]

2. The outbreak of the <u>measles</u> was unexpected.

3. The <u>outbreak</u> of the <u>measles</u> was unexpected. The <u>outbreak</u> occurred yesterday in a school downtown.

4. The emergence of these <u>diseases</u>.

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

5. The <u>elimination</u> of <u>measles</u>.

Adjectives normally denote stative events. Adjectives are to be annotated only when they denote temporally bounded properties or when their presence implies the actual occurrence of another event. When in predicative position, only the adjectival head of the predicative construction must be annotated. As for adjectives in attributive position, they can be marked as events only when their occurrence implies actual occurrences, such as the events leading up to its own existence. Adjectives used as specifiers or mere modifiers of entities (i.e. to restrict the description of an object or person) are not annotated as event.

In this corpus we only annotate events related to the vaccination debate. In order to determine what is a vaccination related event, we looked at the main arguments in favour and against vaccinations, which are presented in websites such as `https://vaccines.procon.org/`. It was not possible to provide the annotators with an exhaustive

list of events that are related to the vaccination debate because the goal is not to annotate events within the vaccination semantic field only, but also events that relate to opinions, beliefs, and attitudes towards vaccinations, legal actions or government campaigns that revolve around vaccinations, economic aspects of vaccinating/or not vaccinating, etc. Annotated events should provide information about the perspectives that entities (persons, institutions, companies) hold about any aspect of vaccinations. Below we list several types of events that are related to the vaccination debate. This can be events:

- Related to: administering and testing vaccines; disease outbreaks; how health organisations/institution, communities, groups, individuals and other entities position themselves towards vaccines; explaining personal experiences with any aspect of vaccines; the achievement that vaccines have brought (vaccines save lives, protect the community, protect future generations); the (adverse) side effects of vaccines: illnesses, symptoms, deaths; the immunity level provided by vaccines; legal aspects of vaccines and policy making; the economic effect of (lack of) vaccination (less illnesses, less expenses for family and society); monitoring the effects of vaccines; commercialising vaccines; vaccines components, ingredients (safe or not, allergies, side effects).

- Referring to actions by persons and institutions in relation to vaccines, the choices that governments do, to personal choices, to the decisions that are made for children, to the right of parents to choose, etc.

- That are mentioned when: reporting discussions related to vaccines; discussing the relation between vaccines and religion or moral attitudes; discussing the personal freedom to choose in relation to vaccines; talking about diseases in relation to vaccines (prevented and (non-)eradicated diseases, vaccine-caused diseases); talking about globalisation and disease spreading.

For this task, a trained student (Annotator A) annotated all the corpus. In order to calculate IAA, one of the authors (Annotator B) annotated 5 documents containing 824 events according to Annotator B. The files to be annotated for IAA where selected randomly. Table 2 shows the results, with an F1 of 88.2.[9]

| Type | P | R | F1 | TP | FP | FN |
|------|------|------|------|------|------|------|
| Event | 88.1 | 88.3 | 88.2 | 726 | 98 | 96 |

Table 2: Annotation of events: IAA in terms of F-scores on 5 files, taking as gold the annotations by Annotator B.

In total, 65,804 vaccine-related events have been annotated by Annotator A, 6,722 of which are unique. 3% (n=2,045) of the annotated events are multi-word expressions, most of

---

[7]`https://github.com/timjogorman/`
`RicherEventDescription/blob/master/`
`guidelines.md`

[8]Events are underlined.

---

[9]In the rest of the paper we use P for precision, R for recall, F1 for F-Score, TP for True Positives, FP for False Positives and FN for False Negatives.

which are disease names such as *Hepatitis B* or *whooping cough*. Table 3 shows the ten most frequently annotated events in the corpus.

| Event | Count | Event | Count |
|---|---|---|---|
| vaccine | 2,998 | risk | 1,085 |
| vaccination | 2,031 | infection | 834 |
| measles | 1,711 | study | 734 |
| disease | 1,502 | report | 606 |
| vaccinate | 1,322 | immunity | 573 |

Table 3: Top 10 most frequent events

The cases of disagreement between annotators are due to several factors. First, cases in which it is not easy to determine whether an event is related to the vaccination debate. Second, some cases in which the annotator clearly forgot to annotate an event. Third, cases that are difficult to annotate, mostly nouns and adjectives.

Some occurrences of event mentions may be complicated to identify as they may correspond/denote or seem to denote more than one entity at the same time. *Vaccination* is the act of administering a vaccine, which means that it is always an event. *Vaccine* is the liquid or preparation that is administered. By default it is not an event. However, it should be annotated as event when it is used in the sense of vaccination through coercion or metonymy, which happens very often. In Examples 6 and 7 vaccine can be interpreted as referring to the act of vaccinating, whereas in Examples 8 and 9 it is less probable:

6. The vaccine was not effective. (Administering the vaccine was not effective)

7. He got a measles vaccine. (To get a vaccine is the same as vaccinating)

8. Mr Holzer voiced concern about Thiomersal (or Thimerosal), a mercury-based preservative used in some vaccines, and about genetically engineered vaccines.

9. They discovered a measles vaccine.

In the following examples annotators disagreed:

10. It was necessary because "of the rising incidence of whooping cough is reported to be related to the vaccine wearing off if given more than ten years earlier.

11. Also, there are relatively benign health implications that are commonly associated with the diseases the vaccine is intended to prevent.

12. The risks that whooping cough, diphtheria and tetanus pose to health are low compared to the potential, serious dangers reported as a result of this vaccine.

Another difficult case are mentions of diseases, which should be annotated as events only if they refer to cases of illness that can be placed in a timeline. In Example 13 "autism" is not an event mention because it does not refer specific cases of autism, but to the illness in general, whereas it is an event in Example 14, because it refers to cases of measles that can be placed on a timeline:

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

14. Some other European countries, including France and Romania, have also seen more measles cases this year.

In the following examples there was disagreement in the annotation of diseases:

15. The report indicates that diphtheria is a bacterial disease.

16. Also, there are relatively benign health implications that are commonly associated with the diseases the vaccine is intended to prevent

17. Also, the media appear to be subtly working to characterize measles as a fatal disease

More cases of nouns that caused disagreement are marked between asterisks in the following examples:

18. The risks that whooping cough, diphtheria and tetanus pose to health are low compared to the potential, serious *dangers* reported as a result of this vaccine.

19. If they should become re-infected, Goldstein writes, subsequent *episodes* (of whooping cough) are generally "quite mild."

20. The *tears* of betrayal following Taylor's post are delicious

21. The psychological *operation* being deployed here is fairly obvious

22. Thanks to the Syrian war and massive *disruptions* of family lives across the Middle East, polio has returned

23. During the past 6 years "aaktionmax" has initiated or supported several public *initiatives* to educate the public

Finally, references to *antivaxers* and *pro-vaccine* caused also disagreements and were annotated by only one annotator:

24. As I noted when his name first came up as a candidate for this position, *antivaxers* weren't going to like it.

25. the FDA Commissioner is the ultimate *pro-vaccine* pharma shill

26. she rejects *antivaccine* pseudoscience explicitly

27. That sure sounds suitably *pro-vaccine* to me

28. They do not cause all the horrible autoimmune and other diseases that *antivaxers* claim

In order to increase IAA for future annotation efforts, we plan to pay special attention to these phenomena in the annotation guidelines.

## 4. Attribution

Attribution is a linguistic phenomenon that consists in "ascribing the ownership of an attitude towards some linguistic material, i.e. the text itself, a portion of it or its semantic content, to an entity" (Prasad et al., 2006; Pareti, 2015). An Attribution Relation (AR) is established when someone (the author of a text, someone mentioned in the text) signals an ownership relation between a third party who expresses an attitude and some text. An AR is composed of three elements (Prabhakaran et al., 2010; Pareti, 2015): a *source* that signals who is providing the information content to a receiver; a *cue* composed of one or more lexical items, which explicitly introduce the AR; and a *content*, which corresponds to the span of text attributed to the source, i.e. what is the source asserting, thinking, believing, promising, etc. The content is the linguistic material which is attributed to the source. Identifying ARs can be conceived as a support task for the extraction of perspectives, since it allows to relate sources to content, and the content can express the target and attitude of a perspective.

ARs can be nested and the content span can be discontinuous. ARs may have three different surface realizations: direct reported speech (signaled by quotation marks), indirect reported speech, and mixed reported speech. Sentence 29 illustrates a case of nested ARs and indirect speech.[10]

29. [While *Trump's team* **denied** [making the specific request that Kennedy referenced], *a spokeswoman* **said** [that Trump is considering forming a commission on autism]], **reported** *CNN and other outlets*.

Some corpora exist with attribution annotations. The Penn Attribution Relations Corpus (PARC) is the first large corpus annotated with ARs to be publicly available. It is an extension of the attribution annotations included in the 2,280 Wall Street Journal articles of the Penn Discourse Treebank (Prasad et al., 2008). Whereas in the PDTB the attribution information is marked only on discourse relations and their arguments, in the PARC corpus ARs are annotated as a phenomenon in its own (Pareti, 2012). The PolNeAR corpus (Newell et al., 2018)[11] is another corpus that contains 1,008 news articles that cover the presidential candidates Hillary Clinton and Donald Trump during the campaign of the 2016 US Presidential Elections. The annotations adhere closely to the annotation scheme of the PARC3 corpus, with some refinements to improve consistency of annotation.

Our annotation guidelines follow closely the guidelines of the PARC corpus. For our task, a trained student (Annotator A) annotated all the corpus. In order to calculate IAA, one of the authors (Annotator B) annotated 25 documents. Table 4 shows the IAA results, with F1 scores above 90.00. In total, Annotator A annotated 4,877 ARs. Out of these, 81% (n=3,929) have exactly one source (Example 30), 18% (n=877) have an implicit source that was not annotated in the text (Example 31), and the remaining ARs have two or more sources (Example 32).

| Type | P | R | F1 | TP | FP | FN |
|---|---|---|---|---|---|---|
| Source | 94.5 | 89.1 | 91.7 | 326 | 19 | 40 |
| Cue | 93.1 | 88.2 | 90.6 | 375 | 28 | 50 |
| Content | 92.3 | 88.9 | 90.6 | 393 | 33 | 49 |

Table 4: Annotation of attribution: IAA in terms of F-scores on 25 documents, taking as gold the annotations by Annotator B.

30. [So far, most of those infected are students or workers at Ohio State], *Brown* **says**.

31. [The Hepatitis B vaccine] is **considered** [one of the safest and most effective vaccines ever made].

32. In fact, **according to** *the FDA* and *the CDC*, [formaldehyde is produced at higher rates by our own metabolic systems and there is no scientific evidence that the low levels of this chemical, mercury or aluminum in vaccines can be harmful].

The majority of ARs (94%, n=4,568) have one cue (as Examples 30–32 above), but there are also some without cues (6%, n=284) or with multiple cues (<1%, n=25) linking the content to the same source (Example 33) or separate sources (Example 34).

33. If *parents' pediatricians* **say** or **imply** [that vaccinating (or vaccinating before a certain age) is dangerous] when it's not, [...].

34. What *we* **suspect**, but *the Department of Health senior advisers* **reject**, is [that if a child has inherited a sensitive immune system then he/she may not cope as well with a vaccine virus].

Out of the total of 4,579 cues, 94% (n=4,303) consisted of a single word, and 6% (n=276) were multi-word expressions (such as *according to*). Table 5 shows the ten most frequently annotated cues.

| Cue | Count | Cue | Count |
|---|---|---|---|
| say | 696 | believe | 181 |
| recommend | 382 | according to | 150 |
| know | 219 | ask | 102 |
| think | 206 | show | 96 |
| tell | 185 | suggest | 96 |

Table 5: Top 10 most frequent attribution cues

## 5. Argumentation: Claims

Numerous models have been developed to address and understand the internal (micro) structure of arguments. A general overview of argumentation schemes is provided in Lippi and Torroni (2016). Stab and Gurevych (2017) created a corpus of 402 argumentative essays from a forum and annotated it with the following argument components: *major claims*, *claims*, and *premises*. They model the microstructure of arguments as a connected tree structure where the *major claim* is the root node which represents

---

[10]The source is marked in italics, the cue in bold and the content between brackets.

[11]Website of the PolNeAR corpus: `https://github.com/networkdynamics/PolNeAR`

the author's standpoint. Al Khatib et al. (2016) constructed a corpus with 100 editorials news websites. The editorials where divided into segments and the segments were assigned one of the following labels: common ground, assumption, testimony, statistics, anecdote, other. Habernal and Gurevych (2017) created a corpus of user-generated Web content collecting documents of different registers. They propose an annotation scheme based on Toulmin's model (Toulmin, 2003), consisting of the following components: claim, grounds, backing, qualifiers, rebuttal, and refutation.

For the annotation of the vaccination corpus we performed two pilot annotation studies with argumentation schemes used in previous annotation tasks (Stab and Gurevych, 2017; Habernal and Gurevych, 2017; Al Khatib et al., 2016). A quantitative and qualitative analysis of the results revealed that it was necessary to simplify the task because the annotation categories were not well defined and, consequently, the IAA was too low (Torsi and Morante, 2018).

Since our final goal is to understand people's attitudes and gain insight in the process of opinion formation, we decided to focus on the core of the argument: the *claims*. Conventionally, an argument is composed at least of two components: a *claim* and a *premise* (Palau and Moens, 2009; Peldszus and Stede, 2013). Because premises are frequently claim-like statements and express the attitude of the source, they were not excluded from the annotation task and they were subsumed in the *claim* component. Therefore, the focus of this task was to identify all claim-like statements. The simplification of the annotation scheme made the annotation task more feasible and less time-consuming. Following the simplified scheme, annotators were better able to agree on fragments of text that are representative of the beliefs that people express when talking about vaccinations. The definition of *claim* chosen was the following: *The claim is the central component of an argument. Claims are sections of text that express the stance of the author. Sometimes, claims are introduced by an explicit source in the text (different from the author). Since they are opinionated statements with respect to the topic, claims are often introduced by stance expressions, such as "In my opinion", "I think that".*

An important requirement is that the *claim* has to be a refutable statement. The following do not qualify as claims: rhetorical questions ("Wouldn't it be better to develop immunity naturally?"), backing ("I am a nurse."), common ground ("Measles can spread through airborne transmission."), statistics ("80% of vaccinated children experience serious side effects."), anecdotes ("I experienced hearing loss after being given the MMR vaccine."), opinions ( "I am against vaccinations.").

Additionally, if the person or entity to whom the *claim* is attributed is different from the author, then annotators should also mark the *source* of the claim.

In future research we plan to study how perspectives are expressed in claims and how we can extract them. Detecting claims should allow to identify the stance of the author or other sources with respect to vaccine related topics. Extracting claims is complementary to extracting attributions, since claims are not necessarily contained in attributions.

We plan to study what is the interrelation between the information expressed in claims and attributions.

The guidelines for the annotation task are released with the annotations. For this task, a trained student (Annotator A) annotated all documents in the corpus resulting in a total of 4,606 claims. In order to calculate IAA, another student (Annotator B), who had less training on the annotation of claims, annotated 50 documents. Table 6 shows the results.

| Type | P | R | F1 | TP | FP | FN |
|---|---|---|---|---|---|---|
| Source | 29.1 | 58.6 | 38.9 | 95 | 231 | 67 |
| Claim | 57.8 | 33.5 | 42.4 | 763 | 556 | 1519 |

Table 6: Annotation of claims: IAA in terms of F-scores on 50 documents, taking as gold the annotations by Annotator A.

The *claim* component achieved 42.4 F-score IAA, which shows that the task remains difficult and needs a better definition. The IAA results for the source of the claim is lower, 38.9 F-score. Annotator B annotated more sources than annotator A, whereas she annotated much less claims than Annotator A. In order to understand the difficulty, we performed an error analysis focusing on the component *claim*. Based on a qualitative error analysis we defined four sources of disagreement:

- *Debatability*, which refers to the degree of debatability that the *claim* needs to express in order to be considered as such. The two annotators followed a different approach when annotating claim-like statements that could be accepted by both the anti-vaccination and the pro-vaccination audiences. Annotator A had the tendency not to annotate such statements.

- *Attributability and commitment*, which refer to the context the *claim* needs to be presented in, specifically looking at whether it can be attributed to a source and how strongly the source needs to commit to it. Annotators exhibited different behaviors in approaching instances where it was unclear to whom the claim-like statement could be attributed.

- *Relatedness to topic*, which refers to deciding whether to annotate or not statements that fit the description of *claim*, but that do not have a direct relation to the topic of vaccination.

- *Granularity and sources*, which refer to different interpretations of the task of assigning sources to the *claim*, resulting in annotations with different degrees of granularity. When explicit sources are mentioned, one annotator decided to annotate *source* and *claim* whereas the other one annotated everything as *claim*. This explains the low agreement for the annotation of sources.

The simplified scheme had several weaknesses. The patterns of disagreement observed reveal the necessity to modify the guidelines for future experiments: (i) more annotated examples of difficult cases should be included; (ii) instructions should be provided on how to tackle instances

where the granularity can be interpreted in different ways; (iii) it should also be stressed that claim-like statements that are not directly related to the topic need to be marked, as they are relevant; and (iv) the guidelines should be more restrictive. For example, the analysis of the errors related to *attributability* lead to the conclusion that *claims* should be attributable to a source and that the source should express a high level of commitment to the claim. A general conclusion is that agreeing on what a *claim* is still remains a difficult endeavor for human annotators. For more details about the annotation see Torsi and Morante (2018).

## 6. Opinions on Persons

In recent years several sets of guidelines have been designed to annotate opinions in texts. Most of these guidelines aim at the identification of sentiment and opinions in social media texts whereas a couple of studies address the identification of opinions in other genres such as news articles ((Wiebe et al., 2005), (Li et al., 2012)). Our annotation guidelines draw on (Wiebe et al., 2005) who present a sentence-level annotation schema for identifying expressions of opinions, beliefs, speculations, sentiments, and other private states in newspaper articles. Their schema focuses on two functional components of private states, i.e. the experiencers holding attitudes and the opinion targets towards which the attitudes are expressed.

For this task we concentrate on the identification of opinion expressions and targets leaving the identification of holders for later work. Additionally, we do not annotate opinions on all topics, but only those on persons, groups of persons and institutions (called Person+ for the remainder of this section). This choice is motivated by the fact that the identification of opinion targets is difficult (see (Sauri et al., 2014)) which we aim to resolve by predefining possible opinion targets in the text. Besides, many different stakeholders participate in a debate like the vaccination debate and their attitude is not only expressed by giving opinions about the topic of vaccination, but also by criticizing and praising each other.

As with most annotations of subjective content (Reidsma and op den Akker, 2008), our guidelines are not specified in extreme detail and the coding relies on the often subjective interpretation of the annotators. There are no fixed rules about how particular words should be annotated and sentences should be interpreted with respect to the contexts in which they appear. The schema we designed accounts for both explicit sentiment (so-called private states), as well as implicit expressions of sentiment (e.g. Examples 35 and 36) and polar facts (e.g. Example 37) (Toprak et al., 2010). Typically, opinions on people include evaluations of their properties (e.g. Example 35) and behaviour (e.g. Example 36 and 37).

35. [A leading conspiracy theorist]$_{expression\_neg}$ is [Andrew Wakefield]$_{target}$

36. [The establishment media]$_{target}$ is [desperately pushing the myth]$_{expression\_neg}$ that vaccines are completely safe and effective

37. [Offit]$_{target}$ has [made millions]$_{expression\_neg}$ inventing a vaccine for Merck

For this task a subset of the corpus is annotated. We selected 210 texts about children's vaccinations thus excluding texts about, for example, the vaccination of animals and travellers. The documents were treated as follows:

- Person+ entities were annotated to create a fixed set of possible opinion targets. 26,996 person+ entities were identified in 210 documents. Table 7 shows the most frequent entities.

- The 210 documents were annotated with opinions by 2 different annotators. They found that only 168 of the documents indeed contained opinions. These 168 documents contain 23,000 Person+ entities and - according to annotator A - 2222 opinions about these Person+ entities.

- 82 out of 168 documents have been double annotated to obtain inter-annotator agreement scores. The results can be seen in Table 8.

| Person+ | Count | Person+ | Count |
|---|---|---|---|
| *personal pronouns* | 7,439 | child(ren) | 650 |
| United States/ US | 444 | CDC | 294 |
| parents | 268 | your child | 154 |
| doctor(s) | 226 | adults | 151 |
| Disneyland | 93 | | |

Table 7: Most frequent Person+ entities

| Type | P | R | F1 | TP | FP | FN |
|---|---|---|---|---|---|---|
| Expressions | 55.8 | 46.0 | 50.5 | 604 | 477 | 707 |

Table 8: Annotation of opinion expressions: IAA in terms of F-scores on 82 documents

Based on an error analysis we saw several sources of disagreement:

- Context-dependentness: Some cases can only be interpreted correctly when the context - ranging from nearby words to the overall stance of the document - is taken into account. In Example 38, for instance, one of the annotators identified *hands out sophisticated, glossy materials* as negative expression about *the group*. This can only be understood from the direct context from which it is clear that the author is annoyed by the fact that this group (of *anti-vaxxers*) is so well-organised.

- The identification of polar facts: polar facts can only be interpreted in a broader context. In Example 39, one of the annotators identified *giving them a MMR shot* as a negative opinion about *the parents* knowing that the overall stance presented in the document was against vaccination. However, annotators easily disagree about these expressions as they do not carry any sentiment by themselves.

- Mutual strengthening of opinions: a major cause of disagreement seems to be the piling up of opinions that

make each other stronger and therefore better to recognize. However, each single expression of such a concatenation can be easily treated differently by the annotators. Example 41 illustrates such a case where two strong negative opinions are given about *Pharma* and where *are a business looking to increase their profits* only can get a negative interpretation because of the earlier opinions.

- Target confusion: in some cases the direct target of the opinion is not a person, but an property related to him or her. In Example 40 one of the annotators considered *revealed as outright fraudulent* as the expression of a negative attitude towards Wakefield whereas the other did not. The disagreement caused by these cases might be avoided by more detailed guidelines.

38. [The group]$_{target}$ [hands out sophisticated, glossy materials]*$_{expression\_neg}$ - in red, white and blue - to lawmakers

39. [their parents]$_{target}$ are planning on [giving them the MMR shot]*$_{expression\_neg}$

40. The study published in 1998 by *[Andrew Wakefield]*$_{target}$ has been [revealed as outright fraudulent]$_{expression\_neg}$.

41. [Pharma]$_{target}$ is [lying to us,]$_{expression\_neg}$ [buying our government officials to FORCE us to pay for chemicals$_{expression\_neg}$ .. they$_{target}$ [are a business looking to increase their profits]*$_{expression\_neg}$

Obviously inter-annotator agreement is low, but most disagreements seem not to stem from systematic differences in the interpretations of the texts by the annotators. Further analysis needs to be done, but we already saw that texts with many opinions seem to have higher agreement (F1 = 0.61) than those with less opinions. This might indicate that identifying opinions in highly opinionated texts is easier than in texts that are not. It might also explain why agreement is low on a corpus with different text genres such as the current corpus.

## 7. Conclusion

In this paper we have presented a corpus of 294 texts from the Internet about the vaccination debate. The texts center around the 2015 measles outbreak linked to the Disneyland Resort in California. The sources are very diverse, including Wikipedia, news, editorials, blogs, (pseudo-)science and a variety of health information dissemination websites from both official state-supported institutions and from independent organizations. The corpus has been compiled with the main goal of having a resource to test methodologies that extract perspectives, in this case about the vaccination debate. By perspectives we mean relations between the source of a statement (i.e. the author or another entity introduced in the text) and a target in that statement (i.e. an entity, event, or (micro-)proposition). This is why we have annotated the corpus with three layers of information that encode aspects of perspectives: attribution, claims, and opinions. Events have also been annotated because we are interested in extracting perspectives on vaccination related events.

The IAA scores reveal that the attribution and event annotation tasks are easier than the claim and opinion tasks. Part of the complexity of the claim task is that we still do not have an accurate definition of what a claim is. As for opinions, it remains difficult to reliably identify expressions that do not have clear opinionated cues and that can only be interpreted in a broader context. The disagreements for the event annotation task were due to several factors: cases in which it is not easy to determine whether an event is related to the vaccination debate, some cases in which the annotator clearly forgot to annotate an event, and cases that are difficult to annotate, mostly of nouns and adjectives. The annotations of *vaccine*, *vaccination* and names of diseases posed challenges. The disagreements for the attribution task were not related to the nature of the task, but mostly to one annotator missing attributions.

In future work we will run more annotation experiments with improved guidelines in order to reach better IAA scores. Based on the lessons learned from this annotation effort, we plan to annotate more files and produce annotations by more annotators, since the corpus in its current state is small and IAA has been calculated on a small number of files. We also would like to annotate files about another controversial topic.

An interesting topic of research is to investigate how the information from the three layers can be integrated. Our future work aims at contextualising the vaccine-related events, ARs, claims and opinions by making explicit how they relate to each other. One way to contextualize information and perspectives is to compare statements in and across texts by means of propositional alignment: which statements entail similar propositions, which express similar perspectives, and which statements present conflicting information. One of our current efforts therefore involves the alignment of propositions containing at least one vaccine-related event and annotating opposition and agreement relations between them. We intend to use these annotations for the evaluation of Natural Language Inference (NLI) or Argument Mining (AM) systems in the context of dispersed online debate.

As it is, this corpus can already be used to study how perspectives are expressed in general and in each of the three annotated layers, as well as to analyze how the layers interact. The corpus can be used also as a test set for systems that process attribution, claims and opinions. Potential applications of the analysis of perspectives on vaccinations would be, for example, detecting beliefs on vaccinations, detecting misinformation about vaccinations to populate ontologies (Amith and Tao, 2018), or training agents (Rosenfeld and Kraus, 2016) to persuade hesitant users to vaccinate.

## 8. Acknowledgements

# 9. Bibliographical References

Al Khatib, K., Wachsmuth, H., Kiesel, J., Hagen, M., and Stein, B. (2016). A news editorial corpus for mining argumentation strategies. In *COLING*, pages 3433–3443.

Amith, M. and Tao, C. (2018). Representing vaccine misinformation using ontologies. *Journal of Biomedical Semantics*, 9:22.

Betsch, C. and Sachse, K. (2012). Dr. jekyll or mr. hyde? how the internet influences vaccination decisions: recent evidence and tentative guidelines for online vaccine communication. *Primary Care: Clinics in Office Practice*, 30(25):3723–3726.

Ceolin, D., Noordegraaf, J., and Aroyo, L. (2016a). Capturing the Ineffable: Collecting, Analysing, and Automating Web Document Quality Assessments. In *Proceedings of the European Knowledge Acquisition Workshop (EKAW)*, pages 83–97.

Ceolin, D., Noordegraaf, J., Aroyo, L., and van Son, C. (2016b). Towards Quality Assessment for Digital Humanities Documents. In *Proceedings of the 8th International ACM Web Science Conference 2016 (WebSci'16)*, Hannover, Germany.

Cybulska, A. and Vossen, P. (2014). Guidelines for ecb+ annotation of events and their coreference. Technical report, Technical report, Technical Report NWR-2014-1, VU University Amsterdam.

Gardner, M., Grus, J., Neumann, M., Tafjord, O., Dasigi, P., Liu, N. F., Peters, M., Schmitz, M., and Zettlemoyer, L. (2018). AllenNLP: A deep semantic natural language processing platform. In *Proceedings of Workshop for NLP Open Source Software (NLP-OSS)*, pages 1–6, Melbourne, Australia, July. Association for Computational Linguistics.

Habernal, I. and Gurevych, I. (2017). Argumentation mining in user-generated web discourse. *Computational Linguistics*, 43(1):125–179.

He, L., Lee, K., Lewis, M., and Zettlemoyer, L. (2017). Deep semantic role labeling: What works and what's next. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 473–483, Vancouver, Canada, July. Association for Computational Linguistics.

ISO, S. W. G., (2008). *ISO DIS 24617-1: 2008 Language resource management - Semantic annotation framework - Part 1: Time and events*. ISO Central Secretariat, Geneva.

Li, H., Cheng, X., Adson, K., Kirshboim, T., and Xu, F. (2012). Annotating opinions in German political news. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, pages 1183–1188, Istanbul, Turkey, May. European Language Resources Association (ELRA).

Lippi, M. and Torroni, P. (2016). Argumentation mining: State of the art and emerging trends. *ACM Transactions on Internet Technology (TOIT)*, 16(2):10.

Maddalena, E., Ceolin, D., and Mizzaro, S. (2018). Multidimensional News Quality: A Comparison of Crowdsourcing and Nichesourcing. In *Proceedings of the 6th International Workshop on News Recommendation and Analytics (INRA 2018)*, Turin, Italy.

Majumder, M. S., Cohn, E. L., Mekaru, S. R., Huston, J. E., and Brownstein, J. S. (2015). Substandard Vaccination Compliance and the 2015 Measles Outbreak. *JAMA Pediatrics*, 169(5):494–495, 05.

Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S. J., and McClosky, D. (2014). The Stanford CoreNLP natural language processing toolkit. In *Association for Computational Linguistics (ACL) System Demonstrations*, pages 55–60.

Mavragani, A. and Ochoa, G. (2018). The Internet and the Anti-Vaccine movement: Tracking the 2017 EU measles outbreak. *Big data and Cognitive Computing*, 2:2.

Mollema, L., Harmsen, I. A., Broekhuizen, E., Clijnk, R., De Melker, H., Paulussen, T., Kok, G., Ruiter, R., and Das, E. (2015). Disease detection or public opinion reflection? Content analysis of tweets, other social media, and online newspapers during the measles outbreak in The Netherlands in 2013. *Journal of medical Internet research*, 17(5).

Newell, E., Margolin, D., and Ruths, D. (2018). An attribution relations corpus for political news. In *Proceedings of the 11th Language Resources and Evaluation Conference*, Miyazaki, Japan, May. European Language Resource Association.

Palau, R. M. and Moens, M.-F. (2009). Argumentation mining: the detection, classification and structure of arguments in text. In *Proceedings of the 12th international conference on artificial intelligence and law*, pages 98–107. ACM.

Palmer, M., Gildea, D., and Kingsbury, P. (2005). The proposition bank: An annotated corpus of semantic roles. *Computational linguistics*, 31(1):71–106.

Pareti, S. (2012). The independent encoding of attribution relations. In *Proceedings of the Eight Joint ACL-ISO Workshop on Interoperable Semantic Annotation (ISA-8)*, Pisa, Italy.

Pareti, S. (2015). *Attribution: A Computational Approach*. Ph.D. thesis, University of Edinburgh.

Peldszus, A. and Stede, M. (2013). From argument diagrams to argumentation mining in texts: A survey. *International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)*, 7(1):1–31.

Phadke, V. K., Bednarczyk, R. A., Salmon, D. A., and Omer, S. B. (2016). Association between vaccine refusal and vaccine-preventable diseases in the united states: a review of measles and pertussis. *Jama*, 315(11):1149–1158.

Prabhakaran, V., Rambow, O., and Diab, M. (2010). Automatic committed belief tagging. In *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, pages 1014–1022, August. Association for Computational Linguistics.

Prasad, R., Dinesh, N., Lee, A., Joshi, A., and Webber, B. (2006). Annotating attribution in the penn discourse treebank. In *Proceedings of the Workshop on Sentiment and Subjectivity in Text*, SST '06, pages 31–38, Strouds-

burg, PA, USA. Association for Computational Linguistics.

Prasad, R., Dinesh, N., Lee, A., Miltsakaki, E., Robaldo, L., Joshi, A., and Webber, B. (2008). The Penn Discourse TreeBank 2.0. *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)*, pages 2961–2968.

Pustejovsky, J., Verhagen, M., Saurí, R., Littman, J., Gaizauskas, R., Katz, G., Mani, I., Knippen, R., and Setzer, A. (2006). TimeBank 1.2. Technical report, Linguistic Data Consortium (LDC), Philadelphia, PA.

Reidsma, D. and op den Akker, H. (2008). Exploiting "subjective" annotations. In R. Artstein, et al., editors, *Coling 2008: Proceedings of the workshop on Human Judgements in Computational Linguistics*, number DTR08-9, pages 8–16. Coling 2008 Organizing Committee, 8.

Rosenfeld, A. and Kraus, S. (2016). Providing arguments in discussions on the basis of the prediction of human argumentative behavior. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 6(4):30.

Sauri, R., Domingo, J., and Badia, T. (2014). A Unifying Opinion Annotation Framework across Genres and in Multiple Languages. In *Proceedings of LREC2014*, pages 11–20.

Stab, C. and Gurevych, I. (2017). Parsing argumentation structures in persuasive essays. *Computational Linguistics*, 43(3):619–659.

Styler, W., Crooks, K., Hamang, M., and O'Gorman, T., (2014). *Richer Event Description (RED) Annotation Guidelines*. University of Colorado at Boulder.

Tonelli, S., Sprugnoli, R., Speranza, M., and Minard, A.-L. (2014). NewsReader Guidelines for Annotation at Document Level. Technical report, Technical report, Technical Report NWR-2014-2, Fondazione Bruno Kessler, Trento.

Toprak, Ã., Jakob, N., and Gurevych, I. (2010). Sentence and expression level annotation of opinions in user-generated discourse. 01.

Torsi, B. and Morante, R. (2018). Annotating claims in the vaccination debate. In *Proceedings of the 5th Workshop on Argument Mining*, pages 47–56, Brussels, Belgium, November. Association for Computational Linguistics.

Toulmin, S. E. (2003). *The uses of argument*. Cambridge University Press.

van Son, C., Caselli, T., Fokkens, A., Maks, I., Morante, R., Aroyo, L., and Vossen, P. (2016). GRaSP: A Multilayered Annotation Scheme for Perspectives. In *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC 2016)*, pages 1177–1184, Portorož, Slovenia.

Wiebe, J., Wilson, T., and Cardie, C. (2005). Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation*, 39.

Wolfe, R. M. and Sharp, L. K. (2002). Anti-vaccinationists past and present. *BMJ: British Medical Journal*, 325(7361):430.