Mining Reasons For And Against Vaccination From Unstructured Data Using Nichesourcing and AI Data Augmentation

Autoría ditelliana: Navajas, Joaquín (*Universidad Torcuato Di Tella. Escuela de Negocios.* Laboratorio de Neurociencia)

Otras autorías: Furman, Damián Ariel (*Universidad de Buenos Aires. CONICET); Junqueras, Juan (Universidad de Buenos Aires);* Gümüslü, Burçe, Deroy, Ophelia; Sulik, Justin *(Ludwig-Maximilians-Universität München);* Altszyler, Edgar (*Universidad de Buenos Aires, Quantit)*

Fecha de publicación: 28/06/2024

Mining Reasons For And Against Vaccination From Unstructured Data Using *Nichesourcing* and AI Data Augmentation

Damián Ariel Furman[1,2], Juan Junqueras[1], Z. Burçe Gümüşlü[3], Edgar Altszyler[1,4], Joaquin Navajas[5], Ophelia Deroy[3], Justin Sulik[3],

[1] Universidad de Buenos Aires, [2] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), [3] Ludwig-Maximilians-Universität München, [4] Quantit, [5] Universidad Torcuato Di Tella

Correspondence: damian.a.furman@gmail.com

Abstract

We present Reasons For and Against Vaccination (RFAV), a dataset for predicting reasons for and against vaccination, and scientific authorities used to justify them, annotated through *nichesourcing* and augmented using GPT4 and GPT3.5-Turbo. We show how it is possible to mine these reasons in non-structured text, under different task definitions, despite the high level of subjectivity involved and explore the impact of artificially augmented data using in-context learning with GPT4 and GPT3.5-Turbo. We publish the dataset and the trained models[1] along with the annotation manual used to train annotators and define the task[2].

1 Introduction

Over the last decades there had been an increase of anti-vaccine propaganda and parents deciding not to vaccinate their children, which have caused outbreaks of diseases that had been previously considered eliminated(Tafuri et al., 2014). The massive development of Internet and communication technologies has provided a mean to facilitate information about vaccines and vaccination campaigns, but also a mean to spread misinformation(Kata, 2010). In this scenario, the development of technologies for automatically recognising what is being said about vaccines can help to rapidly identify new misinformation campaigns and elaborate informed counter-narratives to mitigate the risk they pose.

In this work we present RFAV (Reasons For and Against Vaccination), a dataset with reasons for and against vaccination labeled through *nichesourcing* and expanded using GPT4 and GPT3.5, on websites downloaded from different sources, in English and Spanish. We also trained different language models using this dataset to automatically identify reasons obtaining promising results. Since this task is highly subjective, we include an analysis of difficulties of the annotation process and assess the capabilities of generative LLMs for data augmentation in token classification tasks.

[1] https://huggingface.co/argmining-vaccines

[2] https://github.com/ArgMiningVaccination/RFAV-Dataset

2 Previous Work

Larson et al. (2022) defines Vaccine Hesitancy as "a state of indecision and uncertainty about vaccination before a decision is made (that) represents a time of vulnerability". Wilson and Wiysonge (2020) presented a thorough study concluding that "there is a significant relationship between organization on social media and public doubts of vaccine safety". In this sense, automatic tools "have potential to counter vaccine hesitancy"(Larson and Lin, 2024), as they can help analyze massive online content. Skeppstedt et al. (2018) used topic models to manually code representative arguments about vaccines. (Qorib et al., 2023) applied sentiment analysis for analyzing twitter user's stances about Covid-19 vaccines and reviewed other 14 studies that performed the same task. Torsi and Morante (2018) analyzed three annotation schemes for identifying argument components using a corpus of structured essays and news about vaccination and found that to achieve acceptable IAA they needed to use a simple scheme with only one component that was not strictly argumentative on itself. We follow a similar approach.

3 Corpus creation

To identify relevant web documents, we generated a list of keywords related to vaccination, including complementary/alternative medicine topics as these are associated with vaccine hesitance (Browne et al., 2015). We used SERAPI to conduct Google and Bing searches with those keywords, retrieving URLs for the top 150 hits per search. As this was a scattergun approach, we next sought to boost the proportion of relevant pages. We considered that any web domain reached by at least 10 unique keywords from our list was likely relevant, so we conducted additional SERPAPI searches focusing on those domains, retrieving up to 40 additional URLS per search per domain. Using the Trafilatura python package (Barbaresi, 2021) we parsed the scraped HTML text for each URL, filtering out documents with fewer than 100 words. We used the TextDescriptives python package (Hansen et al., 2023) to excise low-quality sections of text and the Presidio Analyzer to sanitize personal identifying information. This yielded a total of 136934 documents in English and 94361 documents in Spanish. We further filtered these documents using a new list of keywords to preserve only those that were relevant to our purpose of annotation. After filtering, 94398 English documents (69% of the corpus) and 66257 Spanish documents (70% of the corpus) remained.

3.1 Defining the task

All documents were labeled with Reasons for or against vaccination and with Scientific Authorities that might be used to support either a pro or an anti-vaccine stance within the document.

We define a Reason to be anything that can potentially be of interest to a person considering vaccination. They are not necessarily argumentative, though all arguments will be considered reasons. Each example may have zero-to-many reasons and each reason will be also labeled with a 'Stance' value using a Likert scale ranging from 1 to 5, defining the stance that the text has towards vaccination in a broad sense, according to the following descriptions:

1. Strongly against vaccination
2. Weakly against vaccination
3. Ambiguous stance or undetermined
4. Weakly supporting vaccination
5. Strongly supporting vaccination

Strong stances differ from Weak ones because they present themselves as conclusive and make their stance explicit. Weak stances, though relevant when considering vaccination, appear less conclusive and don't have an explicit posture.

We define a Scientific Authority to be any mention or invocation of scientists, publications, scientific, medical or governmental institutions used to provide credibility for potential reasons in the example (either for or against). The link between reasons and scientific authorities does not have to be explicit if it can be inferred that the authority is being cited to provide credibility.

More detailed descriptions of the categories defined and decisions about annotations with examples showing typical cases can be found in the Annotation Manual.

In order to assess how the different categories of stances affect both human and machine performance, we define three tasks related to identification of Reasons: A - a per-word binary classification indicating if a word is part of a reason or not; B - a per-word classification using six categories: 0 for words not belonging to a reason and 1 to 5 to indicate stances; C - a per-word classification using four categories (0 to 3) consisting of a compressed version of the stances that doesn't account for the Weak vs Strong distinction. Considering also the task of predicting Scientific Authorities, this yields 4 different tasks.

3.2 Data annotation

We took a random sample of 1000 documents in English and 1000 examples in Spanish. We removed non-ascii characters and truncated the text to 4000 words, avoiding leaving unfinished sentences when possible. Annotation was performed through *nichesourcing* by six psychology and philosophy advanced college students divided in two teams for English and Spanish. Nichesourcing is "a specific form of outsourcing that harnesses the computational efforts from niche groups of experts rather than the 'faceless crowd'"(Boer et al., 2012). Annotators were asked to carefully review the annotation manual and take a 2 hours course where vaccination-relevant concepts were explained and annotation criteria and examples were discussed.

Each of them then, labeled 400 examples: 100 were common to the three annotators in the same team while the other 300 were exclusive to each individual. This resulted in a total of 1000 examples labeled for each language, with 100 of those labeled three times used for calculating agreement.

Annotation was done using the brat annotation tool (Stenetorp et al., 2012) in three stages. On the first and second stage all members of each team annotated 10 and 30 examples respectively from the common batch and did a pair-review discussing the cases where most disagreement arose. Criteria adopted on these stages was added to the annotation manual. On the third stage, all annotators from each team annotated the last 60 examples from the common batch and then the other 300 examples

ENGLISH	R1	R2	R3	All
Reason	0.50	0.49	0.49	0.49
Compressed stance	0.40	0.45	0.43	0.44
Stance	0.36	0.38	0.36	0.36
Scientific Authority	0.41	0.20	0.51	0.45
SPANISH	R1	R2	R3	All
Reason	0.54	0.50	0.48	0.49
Compressed stance	0.54	0.46	0.47	0.47
Stance	0.36	0.39	0.40	0.39
Scientific Authority	-0.002	0.16	0.31	0.25

Table 1: Cohen's Kappa agreement for English and Spanish. Table shows the average of the agreement between each possible pair of annotators from the three annotators for each language, divided between each round of annotation (1 to 3) and considering all three rounds

from their individual batch of examples.

3.2.1 Agreement

Agreement was calculated, for each language, using Cohen's κ(Cohen, 1960) between all possible pairs of annotators among the three members of the same language team. The reported score is the average of the three agreement values calculated for each combination of the three annotators.

Agreement was calculated in a per-word basis according to the four tasks defined in section 3.1. For "Reason" and "Scientific Authority", agreement is calculated using a binary classification while for "Stance" and "Compressed Stance" is calculated for a multi-label classification with 6 and 4 classes respectively.

Table 1 shows the agreement scores per component and per annotation round.

Based on Cohen's interpretation, Reason, Compressed Stances and Scientific Authority reach a moderate agreement, while Stance shows a fair agreement. For Spanish, Reason and Compressed Stances show a moderate agreement while Stance and Scientific Authority show a Fair agreement (being Stance, very close to moderate). The different values on each round of annotation show that even though the amount of examples was increased progressively, the level of agreement remains except for the case of Scientific Authority, where agreement improves with each round. This lead us to think that pair-reviews helped the annotators reach a better criteria.

Considering that annotation was performed on unstructured documents from different sources not necessarily vaccine-related, we consider this level of agreement to be satisfactory. Agreement is still on the same range of interpretation than (Poudyal et al., 2020), who achieved a Kappa agreement of 0.58 labeling arguments in a corpus of ECHR (European Court of Human Rights) decisions, considering they worked on more argumentatively structured examples. (Furman et al., 2023) labeled argumentative components as a binary classification obtaining an agreement score that ranges from 0.52 to 0.64 depending on the category. Torsi and Morante (2018) report 57% annotator's match ratio on claim detection on debates about vaccination, using a metric ranging from 0 to 1 instead of Cohen's κ that ranges from -1 to 1.

3.2.2 Data Statistics

Figure 1 shows the distribution of words inside recognized reasons that were labeled for each class of Stance. Reasons supporting vaccination (either Strongly or Weakly) constitute 71.59% of the total amount of Reasons labeled on the English dataset and 81,94% on the Spanish dataset, while Reasons against vaccination are 20.76% for English and 13.57% for Spanish. Reasons Strongly Against vaccination are specially scarce in both Spanish and English.

Analyzing the data, we found that many of the documents that seemed to have been scraped from alternative medicine sources didn't mentioned vaccination and were filtered using the keywords as described in 3. We manually reviewed the 100 examples used for agreement calculation and found that most documents from these sources that mentioned vaccines avoided taking explicit stance. Some example of reasons against vaccination found are advertise possible secondary effects (sometimes selling treatment), not enforcing vaccination during Covid pandemic and, from a scientific perspective, also narrowing scope of vaccination campaigns.

3.3 Data augmentation using GPT4 and GPT3.5Turbo

We used OpenAI's GPT4 to annotate 1000 new examples in each language and GPT3.5Turbo to annotate 2900 and 2400 new examples in English and Spanish respectively with Reasons and their Stances, spending U$S600 on GPT4 and U$S65 on GPT3.5-Turbo.

We instructed the model to add [Begin:Reason:*Stance*] at the beginning of a reason and [End:Reason] at the end, where *Stance* stands for a value ranging from 1 to

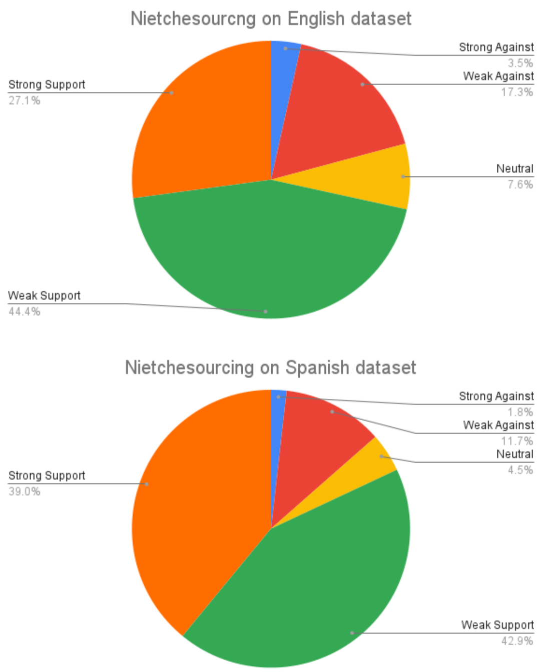

Figure 1: Distribution of labeled words per annotation class on English and Spanish expert annotated dataset

	Examples		Words labeled	
	EN	SP	EN	SP
GPT4	8.3%	9.8%	24%	26%
GPT3.5	12%	14%	24%	24%
Human	44.2%	59.3%	14%	10%

Table 2: Percentage of examples with no Reason labeled (left) and percentage of words that formed part of a Reason (right) in English and Spanish for annotations using GPT4, GPT3.5 and *nichesourcing*

5. The prompt also includes descriptions of the components to be annotated and instructions taken from the annotation manual, providing the model with similar information than human annotators. It also contains a three-shot learner with three labeled examples manually selected to contain reasons with diversity of stances.

While most examples annotated this way respected the proposed format, we found 11 cases in English and 6 cases in Spanish where the end token [End:Reason] was added before any start token. These cases were discarded and replaced with new ones.

Though the prompt instructed the model not to modify in any sense the original example, we noticed that GPT4 and GPT3.5Turbo sometimes introduced some minor changes like adding punctuation symbols at the beginning or at the end of the example, correction of orthography mistakes or syntactic errors or abruptly ending generation though not all words from the original example were processed. We considered that these cases didn't constitute a significant change over the original example and found that the result of the annotation could be used without much problems for training models on all proposed tasks.

3.3.1 Data Statistics

Table 2 shows the percentage of the examples that have no Reason labeled on them and also the percentage of words that are labeled as being part of a Reason, for English and Spanish and for corpus annotated through GPT4, GPT3.5-Turbo and *nichesourcing* (humans). Though values are similar for GPT4 and GPT3.5-Turbo, it can be seen that human annotators labeled proportionally almost half the amount of reasons. Figure 2 shows the distribution of words inside recognized reasons that were labeled for each class of Stance, for GPT4 and GPT3.5-Turbo and for English and Spanish respectively. For GPT4, reasons supporting vaccination in the English dataset (either Strongly or Weakly) constitute 67.33% of the total amount of Reasons while Reasons against vaccination are 19.96%. In the Spanish dataset, reasons supporting vaccination labeled by GPT4 are 68.42% while reasons against vaccination are 22.48%.

For GPT3.5-Turbo, reasons supporting vaccination in the English dataset constitute 47.08% of the total amount of Reasons while Reasons against vaccination are 10.60%. In the Spanish dataset, reasons supporting vaccination labeled by GPT3.5-Turbo are 40.25% while reasons against vaccination are 10.69%.

In all cases, the proportion of Reasons labeled as "Strong Against" and specially the proportion of Reasons labeled as "Neutral" is much higher comparing to human annotators. In particular, for GPT3.5-Turbo, the Neutral class constitutes the majority class by a significant percentage (42.3% for English and 49.1% for Spanish), while the "Strong Support" that constitutes the majority class in all other datasets is greatly diminished in comparison. We manually reviewed 20 examples that were not labeled with reasons by humans and found that GPT4 usually predicted sentences with a positive stance towards medical or scientific procedures in general as a "Support" reason and sentences with a positive stance towards Alternative Medicine related concepts as "Against" disregarding if they were referring to vaccination, while GPT3.5-Turbo usually labeled them as Neutral. Apart from that, we found many annotations that seemed to be reasonable but that differed with the criteria taken by the human annotator.

4 Experiments

Five pre-trained models (two in English, two in Spanish and one Multilingual) were fine-tuned using the English, Spanish and both portions of the corpus respectively, to automatize the tasks defined in 3.1. Datasets were partitioned randomly in three parts for train, development and test, respecting a proportion of 80%, 10% and 10% respectively. We explored three different learning rate values (2e-05, 1e-06 and 2e-06) and kept the model that had the best F1 score on the development partition.

We present a description of the models used:

RoBERTa (Liu et al., 2019) is a transformer English language model based on BERT (Devlin et al., 2019) that established a new state-of-the-art for 4 out of 9 GLUE tasks and matched state-of-the-art on other 2.

LongFormer (Beltagy et al., 2020) is a transformer based English language model specially designed for processing long documents. It is initialized using the weights of RoBERTa and then further pre-trained again on a corpus of 100K long documents to induce learning of long-range dependencies.

XLM Roberta (Conneau et al., 2020) is a transformer model based on Roberta architecture but trained with 2.5TB of data containing 100 different languages.

BETO (Cañete et al., 2020) is a transformer based on BERT but trained from scratch from a big compilation of Spanish unannotated corpus from 15 different sources.

SpanBERTa (Tran, 2020) is a model developed by SkimAI based on RoBERTa's architecture but trained from scratch with 18GB of Spanish data from a big corpus compiled from different sources.

4.1 Training models with human annotated data

4.1.1 Evaluation

For each model, we report F1, Precision and Recall scores over predictions on test dataset. For the two multi-label tasks of Stance recognition we

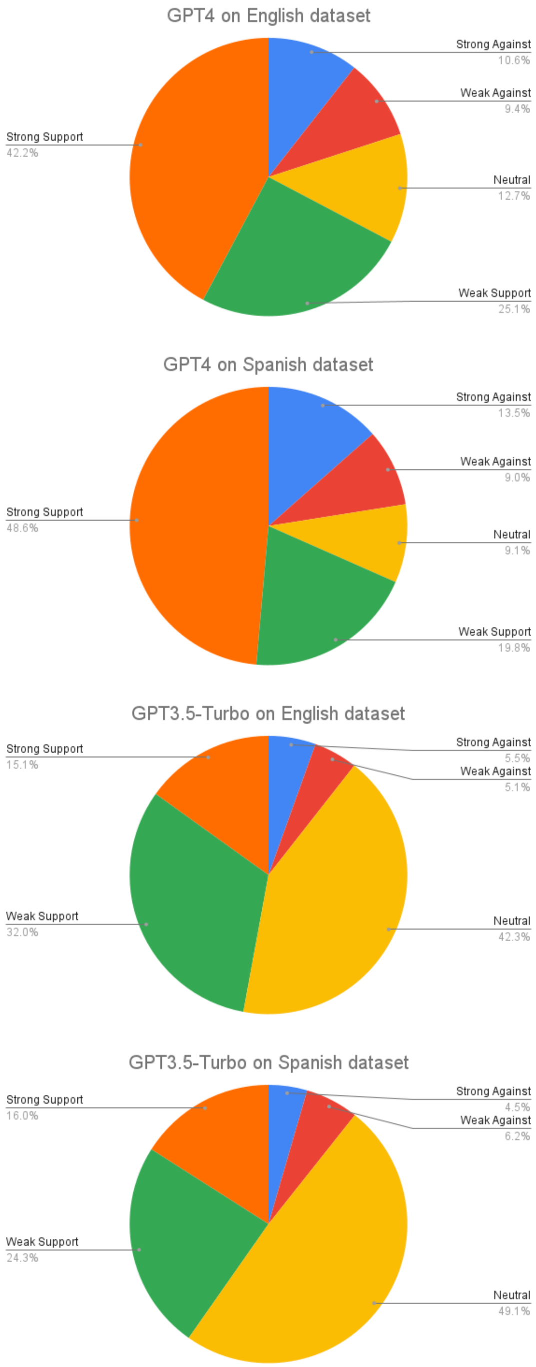

Figure 2: Distribution of labeled words per annotation class by GPT4 and GPT3.5-Turbno on English and Spanish datasets

also report F1 scores per category. Scores must be interpreted considering the subjective nature of the task so for the sake of comparison, we took the examples labeled by the three annotators (used to calculate agreement) and calculated F1 scores of all possible pairs considering one to be the ground truth and the other, a human predictor trying to mimic the other annotator (and vice-versa). We use this score as an indicative measure of a machine predictor's performance compared to a human, per each category. We report the average of the six F1 scores calculated, the worst and the best scores, for English and Spanish.

4.1.2 Results

Table 3 shows the results obtained by all the classifiers trained for the task of Automatic Recognition of Reasons. For all models except BETO, performance was close or even slightly above Human annotators. Longformer performed best with an F1 score of 0.64. BETO performed more than 10 points below the other Spanish model, SpanBERTa. Table 4 shows the results obtained by all the classifiers trained for the task of Automatic Recognition of cites of Scientific Authority. RoBERTa classifier obtained a 0.43 F1 score, above average human performance. Longformer classifier obtained a much lower score, demonstrating that long-range dependencies are not important for this task. For Spanish, Human performance is much lower, which corresponds to the lower agreement values shown in table 1 so both models are above their performance. In this case, BETO obtained the higher score.Table 5 shows the results obtained by all the classifiers trained on the task of Automatic Recognition of Stances, predicting both if a word belongs to a reason and its Stance value. This is a difficult task because it involves classification using six labels with highly unbalanced distribution (see figure 1). Table 7 shows the F1 score per class for each model. English models achieve an acceptable performance for recognizing Support stances while they have no performance at all for Strong Against and Neutrals, both classes that were least frequents on the dataset. Only Longformer showed a better performance for Weak Against.
Table 6 shows the results obtained by all the classifiers trained in the task of Automatic Recognition of Compressed Stances. Results for all models rise between .11 and .14, a significant improvement, compared to the non-compressed version. This leads to interpret that a great amount of "mistakes" considered in the scoring of the models where because of difficulties at recognizing Strong vs Weak stances but not at recognizing Against vs Pro stances.
Unlike previous experiments, here the best model's performance (Longformer for English and SpanBERTa for Spanish) is .08 and .07 below human performance, respectively.
Table 8 shows the F1 scores per class for this tasks. Spanish and Multilingual models still show no performance for Against class. Roberta, however, improved its performance significantly compared to both Against classes considered separately.
Pro class shows an improvement compared to both Pro classes from the the task of Stance recognition yielding a good performance, close to the binary task of recognizing reasons.
A manual examination of examples labeled by humans and by automatic models can be found on appendix section B.1.

Model	F1	Pr	Rec
Roberta (EN)	0.56	0.69	0.47
Longformer (EN)	**0.64**	0.58	0.72
XLM-Roberta (Multi)	0.59	0.66	0.53
SpanBERTa (SP)	0.58	0.47	0.77
BETO (SP)	0.47	0.50	0.44
Avg Human English	**0.56**	0.56	0.56
Best Human English	0.58	0.64	0.53
Worst Human English	0.52	0.5	0.54
Avg Human Spanish	**0.53**	0.54	0.54
Best Human Spanish	0.57	0.71	0.47
Worst Human Spanish	0.49	0.46	0.54

Table 3: F1, Precision and Recall scores of different models for the task of predicting reasons within an example

Model	F1	Pr	Rec
Roberta (EN)	**0.43**	0.38	0.51
Longformer (EN)	0.29	0.68	0.19
XLM-Roberta (Multi)	0.25	0.49	0.18
SpanBERTa (SP)	0.27	0.46	0.20
BETO (SP)	0.36	0.38	0.33
Avg Human English	**0.42**	0.5	0.5
Best Human English	0.45	0.7	0.3
Worst Human English	0.38	0.25	0.83
Avg Human Spanish	**0.25**	0.28	0.28
Best Human Spanish	0.4	0.53	0.32
Worst Human Spanish	0.17	0.32	0.12

Table 4: F1, Precision and Recall scores of different models for the task of predicting scientific authorities

Model	F1	Pr	Rec
Roberta (EN)	0.28	0.33	0.28
Longformer (EN)	**0.31**	0.35	0.30
XLM-Roberta (Multi)	0.2	0.26	0.19
SpanBERTa (SP)	0.26	0.26	0.26
BETO (SP)	0.24	0.29	0.23
Average Human English	**0.36**	0.38	0.38
Best Human English	0.54	0.53	0.54
Worst Human English	0.22	0.21	0.29
Average Human Spanish	**0.31**	0.33	0.33
Best Human Spanish	0.32	0.4	0.32
Worst Human Spanish	0.28	0.31	0.26

Table 5: F1, Precision and Recall scores of different models for the task of predicting stances

Model	F1	Pr	Rec
Roberta (EN)	0.43	0.48	0.41
Longformer (EN)	**0.43**	0.48	0.4
XLM-Roberta (Multi)	0.36	0.35	0.38
SpanBERTa (SP)	0.36	0.34	0.39
BETO (SP)	0.35	0.35	0.34
Average Human English	**0.51**	0.51	0.51
Best Human English	0.54	0.53	0.56
Worst Human English	0.48	0.49	0.47
Average Human Spanish	**0.43**	0.44	0.44
Best Human Spanish	0.45	0.49	0.42
Worst Human Spanish	0.41	0.45	0.39

Table 6: F1, Precision and Recall scores of different models for the task of predicting a reduced set of stances (three instead of five)

	Against		Neu	Support	
Model	Str	Wk		Wk	Str
Roberta (EN)	.0	.05	.0	.26	.45
Longformer (EN)	.0	.27	.0	.20	.46
XLM-Roberta (Multi)	.0	.0	.0	.14	.14
SpanBERTa (SP)	.0	.0	.0	.31	.31
BETO (SP)	.0	.0	.0	.21	.33

Table 7: F1 Scores for the task of detecting stances per each class: Strong Against, Weak Against, Neutral, Weak Support and Strong Support

Model	Against	Neutral	Pro
Roberta (EN)	.23	.0	.56
Longformer (EN)	0.27	.0	.52
XLM-Roberta (EN)	.0	.0	.54
SpanBERTa (SP)	0.01	.0	.50
BETO (SP)	0.00	.0	.45

Table 8: F1 Scores per class for the task of detecting a compressed version of stances

	Hum + GPT4			All		
Model	F1	Pr	Rec	F1	Pr	Rec
RoBERTa	.45	.70	.33	.31	.71	.20
Longformer	.39	.78	.26	.19	.83	.11
XLM-Roberta	.52	.54	.50	.10	.78	.05
SpanBERTa	.48	.51	.46	.03	.73	.02
BETO	.43	.61	.33	.20	.83	.11

Table 9: Results of models trained with both Human + GPT4 and Human + GPT4 + GPT3.5 (All corpora) for predicting Reasons

	Hum + GPT4			All		
Model	F1	Pr	Rec	F1	Pr	Rec
RoBERTa	.27	.28	.29	.21	.23	.22
Longformer	.22	.21	.23	.15	.14	.17
XLM-Roberta	.27	.26	.28	.18	.24	.18
SpanBERTa	.23	.22	.24	.21	.26	.22
BETO	.20	.34	.19	.22	.48	.20

Table 10: Results of models trained with both Human + GPT4 and Human + GPT4 + GPT3.5 (All corpora) for predicting Stances.

4.2 Training using augmented data

Table 9 shows the results of predictions of Reasons done by models trained by combining the dataset labeled through *nichesourcing* with only GPT4 annotated dataset and with both GPT4 and GPT3.5-Turbo annotated datasets. Models were tested against the same test partition used for experiments in section 4.1.2. It can be seen that by combining the Human annotated training partition with these datasets, the overall performance decreased. The more data we use for training the worse result we get. The models whose performance decreased the most are those who had a better performance when training only with the Human annotated corpus. Tables 10 and 11 show the results of training models with the same combination of datasets for predicting Stances and the Compressed Stances respectively. Again, we observe a decrease in model's performances but much smaller than when analysing Reasons.

In order to gain insights for analyzing these re-

	Hum + GPT4			All		
Model	F1	Pr	Rec	F1	Pr	Rec
RoBERTa	.42	.49	.39	.32	.59	.32
Longformer	.30	.36	.29	.36	.52	.34
XLM-Roberta	.38	.43	.36	.23	.40	.25
SpanBERTa	.38	.40	.37	.31	.37	.29
BETO	.35	.52	.33	.35	.58	.32

Table 11: Results of models trained with both Human + GPT4 and Human + GPT4 + GPT3.5 (All corpora) for predicting Compressed Stances

Model	Hum	Hum+ GPT4	All
RoBERTa	12.5%	8.6%	5.1%
Longformer	22.6%	6.1%	2.4%
XLM-Roberta	14.6%	12%	1.2%
SpanBERTa	21.5%	3.9%	0.3%
BETO	11.5%	7.2%	1.8%

Table 12: Percentage of words labeled by predictor as Reasons, for predictor trained with Human, Human + GPT4 and Human + GPT4 + GPT3.5Turbo annotated data. This is, the percentage of True positives + False positives over the whole dataset.

sults, we evaluated the performance of GPT4 and GPT3.5-Turbo against the test dataset using human annotations as gold standard. From 100 examples, 71 in English and 70 in Spanish were unchanged after adjusting the output with a postprocessing script that removes possible additions by GPT. The rest of the examples were discarded.

Tables 13 and 14 show F1, Precision and Recall scores for Automatic Recognition of Reasons, Stances and Compressed Stances for annotations done with GPT4 and compare those values to the ones obtained by the best model and human evaluation from section 4.1.2.

These results seem to suggest that GPT models with fewshot learners don't perform as well as smaller open-source models finetuned with high quality data labeled by experts, or at least, they are not able to absorb the subjective criteria defined through the annotation process only by prompting and in-context learning.

Table 12 shows the percentage of words that were labeled as being part of a Reason by each model. This is, the percentage of the annotated data that is either a True or a False positive. We can observe that the more data is used for training, the more conservative the model trained with that data becomes when predicting on the test dataset. This may seem contradictory on a first inspection given that the augmented data have almost twice as much positive labels than Human annotated examples (see section 3.3.1). Our hypothesis is that the combination of datasets with different annotation criteria affects negatively the models predictive capacity making them more conservative.

5 Conclusions

In this work we present a protocol for annotating reasons with a stance towards vaccination and a dataset of 1000 examples in English and 1000 examples in Spanish annotated by six persons through

Component	F1	Pr	Rec	Best F1	Hum F1
Reasons	0.43	0.44	0.44	0.64	0.56
Stances	0.26	0.26	0.33	0.31	0.36
Compressed	0.39	0.40	0.40	0.43	0.51

Table 13: Performance of GPT4 on the English test dataset for detecting reasons, stances and compressed stances, compared with the best model and human F1 scores for reference

Component	F1	Pr	Rec	Best F1	Hum F1
Reasons	0.40	0.44	0.44	0.58	0.53
Stances	0.23	0.27	0.27	0.26	0.31
Compressed	0.35	0.34	0.38	0.36	0.43

Table 14: Performance of GPT4 on the Spanish test dataset for detecting reasons, stances and compressed stances, compared with the best model and human F1 scores for reference

Nichesourcing and 3900 examples in English and 3400 examples in Spanish annotated using GPT4 and GPT3.5-Turbo with a fewshot learner and a short synthesis of the annotation manual on the prompt. We release the dataset and the finetuned models for the free use of the scientific community. Despite the highly subjective nature of the task, we achieved an acceptable IAA thanks to an iterative annotation process where annotation criteria and examples were discussed. Annotation manual registering this process is also released. Experiments show that the annotation process can be reproduced automatically with satisfactory results considering the level of subjectivity of the correspondent task measured using Cohen's Kappa and the F1 scores of all combinations of annotators, with some room for improvement on the tasks of detecting Stances, particularly for the Against and Neutral classes. When augmenting the human annotated corpus using annotations performed by GPT4 and GPT3.5-Turbo performance decreased, specially for the task of automatically identifying Reasons. Manual inspection of the augmented data revealed that the annotations made by GPT models were not senseless but rather followed a different criteria than human experts, tending to consider a wider range of subjects to be vaccine related (leading to annotate approximately 80% more examples and twice the amount of words). We conclude that GPT models were not able to reproduce the annotation criteria of human annotators only by incorporating a reduced version of the annotation manual and three examples on the prompt.

6 Limitations

In the following section we acknowledge some limitations found in our work.

Experiment results show that data imbalance of the annotated corpus directly affects the predictive capabilities of the models. Results from tables 7 and 5 show that performance for the majority class ("Support") is much higher, while performance for the "Against" or "Neutral" classes is lower. This is related to the fact that these categories are scarce in the annotated dataset, as can be observed on figure 1. This data imbalance is a reflection of the proportion of online content supporting and attacking vaccination, being the first one much common than the second. Therefore, the only way to augment the sample of minority classes without artificially altering the distribution of classes is to label more examples, which is costly. This limits a possible use for the tool: to automatically recognize what is being said against vaccination in order to help elaborate adequate responses. More work needs to be done in order to improve model performance on minority classes.

Our strategy of data augmentation using generative models like GPT4 and GPT3.5-Turbo was based on providing them a summary of the annotation manual within the prompt and using in-context learning to make them learn the annotation criteria. However, annotation produced by these models followed a different criteria than human annotators, tending to consider a wider range of statements to be vaccine-related, therefore producing a different distribution of classes, which can be observed in figure 2. While GPT4 tended to consider any statement that was science-related to be of the Pro class and any statement relative to alternative medicine to be of the Against class, GPT3.5 tended to annotate a lot of vaccine unrelated content as a Neutral Reason. We believe that this difference in annotation criteria made the models that were finetuned using this data combined with human annotations to become even more conservative in their labelling, specially over the minority classes, thus achieving lower results.

7 Ethical Considerations

Though this tool is intended to be used to fight misinformation campaigns causing vaccine hesitancy and possibly, outbreaks of preventable diseases, it could also be used as a tool to mine Reasons supporting vaccination in order to orient misinformation campaigns to target most commonly used reasons supporting vaccination. We acknowledge this possible misuse of our tool but we also reason that contrasting arguments, facts and information should help people to take more informed and rational decisions in the end.

Though one of our goals is to fight misinformation to help prevent outbreaks of preventable diseases, we also want to acknowledge that not all reasons against vaccination are necessarily misinformation. Example 15 in appendix shows a reason against vaccination of immunosuppressed patients against COVID-19 based on lack of testing and the possibility to wait given that there was little cases in that country at that time. We found a significant amount of examples like this one, where reasons for not getting vaccinated were presented not against vaccination in general, but against a particular vaccine or vaccination campaign and they were presented with a scientific base. The dataset along with the trained models presented in this work are meant to help to automatically identify what is being said about vaccines and vaccination. It must be used with caution and critical thinking.

References

A. Barbaresi. 2021. Trafilatura: A web scraping library and command-line tool for text discovery and extraction.

Iz Beltagy, Matthew E. Peters, and Arman Cohan. 2020. Longformer: The long-document transformer. *arXiv:2004.05150*.

Victor Boer, Michiel Hildebrand, Lora Aroyo, Pieter De Leenheer, Chris Dijkshoorn, Binyam Tesfa, and Guus Schreiber. 2012. Nichesourcing: Harnessing the power of crowds of experts. volume 7603, pages 16–20.

M. Browne, P. Thomson, M. J. Rockloff, and G. Pennycook. 2015. Going against the herd: psychological and cultural factors underlying the 'vaccination confidence gap'.

José Cañete, Gabriel Chaperon, Rodrigo Fuentes, Jou-Hui Ho, Hojin Kang, and Jorge Pérez. 2020. Spanish pre-trained bert model and evaluation data. In *PML4DC at ICLR 2020*.

J. Cohen. 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1):37.

Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. 2020. Unsupervised cross-lingual representation learning at scale. *Preprint*, arXiv:1911.02116.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Damián Furman, Pablo Torres, José A. Rodríguez, Diego Letzen, Vanina Martínez, and Laura Alonso Alemany. 2023. Which argumentative aspects of hate speech in social media can be reliably identified? *Preprint*, arXiv:2306.02978.

L. Hansen, L. R. Olsen, and K. Enevoldsen. 2023. Textdescriptives: A python package for calculating a large variety of metrics from text.

Anna Kata. 2010. A postmodern pandora's box: Anti-vaccination misinformation on the internet. *Vaccine*, 28(7):1709–1716.

Heidi J. Larson, Emmanuela Gakidou, and Christopher J.L. Murray. 2022. The vaccine-hesitant moment. *New England Journal of Medicine*, 387(1):58–65.

Heidi J Larson and Leesa Lin. 2024. Generative artificial intelligence can have a role in combating vaccine hesitancy. *BMJ*, 384.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint*.

Prakash Poudyal, Jaromir Savelka, Aagje Ieven, Marie Francine Moens, Teresa Goncalves, and Paulo Quaresma. 2020. ECHR: Legal corpus for argument mining. In *Proceedings of the 7th Workshop on Argument Mining*, pages 67–75, Online. Association for Computational Linguistics.

Miftahul Qorib, Timothy Oladunni, Max Denis, Esther Ososanya, and Paul Cotae. 2023. Covid-19 vaccine hesitancy: Text mining, sentiment analysis and machine learning on covid-19 vaccination twitter dataset. *Expert Systems with Applications*, 212:118715.

Maria Skeppstedt, Andreas Kerren, and Manfred Stede. 2018. Vaccine hesitancy in discussion forums: Computer-assisted argument mining with topic models. *Studies in health technology and informatics*, 247:366–370.

Pontus Stenetorp, Sampo Pyysalo, Goran Topić, Tomoko Ohta, Sophia Ananiadou, and Jun'ichi Tsujii. 2012. brat: a web-based tool for NLP-assisted text annotation. In *Proceedings of the Demonstrations at the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pages 102–107, Avignon, France. Association for Computational Linguistics.

S. Tafuri, M.S. Gallone, M.G. Cappelli, D. Martinelli, R. Prato, and C. Germinario. 2014. Addressing the anti-vaccination movement and the role of hcws. *Vaccine*, 32(38):4860–4865. Vaccine-preventable Diseases and Vaccinations Among Health-care Workers.

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

Chris Tran. 2020. https://github.com/chriskhanhtran/spanish-bert.

Steven Lloyd Wilson and Charles Wiysonge. 2020. Social media and vaccine hesitancy. *BMJ Global Health*, 5(10).

```
Extract reasons either supporting or opposing
vaccination, and link them to the corresponding
stance values.
• A Reason potentially or hypothetically appeals
to someone considering vaccination.
• They must be something relevant to someone
hypothetically considering getting or not
getting vaccinated.
• Examples can have zero or many reasons.
• Reasons have a number indicating their stance
towards vaccination.
• The token [Reason:begin:1] indicates a reason
that is strongly against vaccination.
• The token [Reason:begin:2] indicates a reason
that is weakly against vaccination, this means,
that it highlights negative aspects associated
with vaccination without explicitly taking a
stance against it.
• The token [Reason:begin:3] indicates a reason
that have a neutral stance towards vaccination or
which stance can not be inferred.
• The token [Reason:begin:4] indicates a reason
weakly supporting vaccination. This means that it
provides positive aspects of vaccination (like
"they are free" or "they are accessible") without
explicitly taking a stance.
• The token [Reason:begin:5] indicates a reason
strongly supporting vaccination. It associates
vaccines explicitly with good qualities and
positive concepts.
• Do not remove the links from the original non
annotated text, keep them in plain text,
respecting the original format.
• An identified reason should be marked with the
special tokens "[Reason:begin:stanceValue]"
before the first word of the reason and
"[Reason:end]" at the end of the reason.
• stanceValue is an integer ranging from 1 to 5.
• The output should contain exactly the same text
as the input only adding the special tokens when
appropriate.
```

Figure 3: Template used for generating prompts for annotation using GPT4 and GPT3.5. The final version of the prompt included three non-annotated examples linked to their correspondent annotations

A Prompt used for data augmentation

Figure 3 shows the template used for generating the prompts. While this was the same for both languages, examples used for in-context learning were selected to match the same language as the example being annotated.

B Example Appendix

B.1 Manual examination of predicted examples

We randomly selected 3 examples from the test set and generated predictions for each of them using GPT4 and the best performing models from each category. Figures 4, 5, 6, 7, 8 and 9 show Example 105 from test dataset as annotated by a human annotator, a finetuned Longformer predicting exclusively Reasons, a finetuned Longformer predicting Reasons with their Stance, a finetuned Longformer predicting Reasons with the Compressed Stances, a finetuned Roberta predicting Scientific Authorities, and GPT4 respectively. It can be observed that there is general agreement about the paragraph starting with "We are ready for a healthier tomorrow...", with the exception of the model predicting Stances, which left the first part of it unannotated. Human annotator considered the first two sentences to be valid reasons and GPT4 considered also the other next paragraph, but none of the pretrained models labeled them. The Roberta model predicting Scientific Authorities labeled "Western Winsconsin Health", though Human annotator didn't.

Figures 10, 11, 12 and 13 show Example 114 from test dataset as annotated, also, by a human annotator, a model predicting exclusively Reasons, a model predicting Reasons with their Stance and a model predicting Scientific Authorities, respectively. In this case, the model predicting Scientific Authorities matched exactly with Human annotation. For the model predicting Reasons, there are minor discrepancies regarding if the phrases starting with "Learn more about..." should be considered a reason or not but mostly matches with Human annotation. When inspecting Stance predictions we observe a higher level of discrepancies regarding what is a reason or not and also about the Stance value though the difference is only between Strong and Weak Support. This example was not labeled by GPT4 because it produced a result that was not correctly labeled (see section 4.2)

In both examples we found that the Longformer model predicting Stances doesn't label the full extent of a sentence, something that was clearly stated on the annotation protocol and that prevailed on most Human Annotations and also mostly on models predicting only Reasons.

Figures 16, 17, 18, 19, 20 and 21 show Example 784 from test dataset as annotated by a human annotator, a finetuned Longformer predicting exclusively Reasons, a finetuned Longformer predicting Reasons with their Stance, a finetuned Longformer predicting Reasons with the Compressed Stances, a finetuned Roberta predicting Scientific Authorities, and GPT4 respectively. Model predicting only Reasons shows a high level of matching with Human annotations, though it predicted one extra sentence on the first paragraph. The paragraph starting with "Currently, Sanofi's..." was only partially labeled by the Human annotator although the annotation manual clearly states that whenever it was possible, whole sentences should be labeled, just like the Longformer model did. Model predicting Stances didn't make any predictions and thus found no Reasons. However, model predicting the compressed version of the Stances did find a Reason which partially matched one of the Reasons labeled by the Human annotator. No Scientific Authority was labeled on this example neither the model predicted one, so in that sense, they match. Lastly, GPT4 labeled the second sentence and not the first, like the Human annotator did. It labeled the whole paragraph starting with "Currently..." just like the annotation manual says, and also labeled the last paragrapgh matching the annotation made by the Human expert.

Lastly, figure 14 shows example 160 from the test dataset which was manually selected because it has a clear stance against vaccination. In this case, no finetuned model predicted any Reasons or Scientific Authorities. GPT4's prediction, on the other hand, matched exactly with Human annotator.

A Healthier Tomorrow? - With the Vaccine - YES! **(5)**
A Healthier Tomorrow? - With the Vaccine - YES! **(5)**
Imagine a Healthier Tomorrow By Kathleen Findlay, MD, MPH - Family Practice and Integrative Medicine Physician Our vision at Western Wisconsin Health is "imagine a healthier tomorrow." These days, we all spend a lot of time imagining a healthier tomorrow-namely one without COVID-19.
We imagine the day when we can go out without wearing a mask, when we can once again meet our friends for a night out, or when we can enjoy our kids' sporting events with their grandparents.
The damage inflicted from lockdowns is felt from our bank accounts to our spirits.
We are ready for that healthier tomorrow, and the good news is that we can see that day sooner thanks to the groundbreaking development of the COVID-19 vaccine **(5)**. Through global cooperation of scientists, researchers, and medical experts, an extraordinarily effective and safe vaccine is now available **(5)**.
The speed of development was only possible due to sharing of research on a scale we have never seen before.Scientists had a head start thanks to prior research done on SARS and MERS.

Figure 4: Example 105 from test dataset labeled through nichesourcing with Reasons, Stances and Scientific Authorities

A Healthier Tomorrow? - With the Vaccine - YES!
A Healthier Tomorrow? - With the Vaccine - YES!
Imagine a Healthier Tomorrow By Kathleen Findlay, MD, MPH - Family Practice and Integrative Medicine Physician Our vision at Western Wisconsin Health is "imagine a healthier tomorrow." These days, we all spend a lot of time imagining a healthier tomorrow-namely one without COVID-19.
We imagine the day when we can go out without wearing a mask, when we can once again meet our friends for a night out, or when we can enjoy our kids' sporting events with their grandparents.
The damage inflicted from lockdowns is felt from our bank accounts to our spirits. We are ready for that healthier tomorrow, and the good news is that we can see that day sooner thanks to the groundbreaking development of the COVID-19 vaccine. Through global cooperation of scientists, researchers, and medical experts, an extraordinarily effective and safe vaccine is now available.
The speed of development was only possible due to sharing of research on a scale we have never seen before. Scientists had a head start thanks to prior research done on SARS and MERS.

Figure 5: Example 105 from test dataset labeled by our finetuned Longformer only with Reasons

A Healthier Tomorrow? - With the Vaccine - YES!
A Healthier Tomorrow? - With the Vaccine - YES!
Imagine a Healthier Tomorrow By Kathleen Findlay, MD, MPH - Family Practice and Integrative Medicine Physician Our vision at Western Wisconsin Health is "imagine a healthier tomorrow." These days, we all spend a lot of time imagining a healthier tomorrow-namely one without COVID-19.
We imagine the day when we can go out without wearing a mask, when we can once again meet our friends for a night out, or when we can enjoy our kids' sporting events with their grandparents.
The damage inflicted from lockdowns is felt from our bank accounts to our spirits.
We are ready for that healthier tomorrow, and the good news is that we can see that day sooner thanks to the groundbreaking development of the COVID-19 vaccine **(4)**. Through global cooperation of scientists, researchers, and medical experts, an extraordinarily effective and safe vaccine is now available **(5)**.
The speed of development was only possible due to sharing of research on a scale we have never seen before.Scientists had a head start thanks to prior research done on SARS and MERS.

Figure 6: Example 105 from test dataset labeled by our finetuned Longformer with Reasons and their Stance

A Healthier Tomorrow? - With the Vaccine - YES!
A Healthier Tomorrow? - With the Vaccine - YES!
Imagine a Healthier Tomorrow By Kathleen Findlay, MD, MPH - Family Practice and Integrative Medicine Physician Our vision at Western Wisconsin Health is "imagine a healthier tomorrow." These days, we all spend a lot of time imagining a healthier tomorrow-namely one without COVID-19.
We imagine the day when we can go out without wearing a mask, when we can once again meet our friends for a night out, or when we can enjoy our kids' sporting events with their grandparents.
The damage inflicted from lockdowns is felt from our bank accounts to our spirits. We are ready for that healthier tomorrow, and the good news is that we can see that day sooner thanks to the groundbreaking development of the COVID-19 vaccine. Through global cooperation of scientists, researchers, and medical experts, an extraordinarily effective and safe vaccine is now available **(3)**.
The speed of development was only possible due to sharing of research on a scale we have never seen before. Scientists had a head start thanks to prior research done on SARS and MERS.

Figure 7: Example 105 from test dataset labeled by our finetuned Longformer with Reasons and the Compressed version of Stances

A Healthier Tomorrow? - With the Vaccine - YES!
A Healthier Tomorrow? - With the Vaccine - YES!
Imagine a Healthier Tomorrow By Kathleen Findlay, MD, MPH - Family Practice and Integrative Medicine Physician Our vision at Western Wisconsin Health is "imagine a healthier tomorrow." These days, we all spend a lot of time imagining a healthier tomorrow-namely one without COVID-19.
We imagine the day when we can go out without wearing a mask, when we can once again meet our friends for a night out, or when we can enjoy our kids' sporting events with their grandparents.
The damage inflicted from lockdowns is felt from our bank accounts to our spirits.
We are ready for that healthier tomorrow, and the good news is that we can see that day sooner thanks to the groundbreaking development of the COVID-19 vaccine. Through global cooperation of scientists, researchers, and medical experts, an extraordinarily effective and safe vaccine is now available.
The speed of development was only possible due to sharing of research on a scale we have never seen before.Scientists had a head start thanks to prior research done on SARS and MERS.

Figure 8: Example 105 from test dataset labeled by our finetuned Roberta-base with Scientific Authorities

A Healthier Tomorrow? - With the Vaccine - YES!
A Healthier Tomorrow? - With the Vaccine - YES!
Imagine a Healthier Tomorrow By Kathleen Findlay, MD, MPH - Family Practice and Integrative Medicine Physician Our vision at Western Wisconsin Health is "imagine a healthier tomorrow." These days, we all spend a lot of time imagining a healthier tomorrow-namely one without COVID-19 **(5)**.
We imagine the day when we can go out without wearing a mask, when we can once again meet our friends for a night out, or when we can enjoy our kids' sporting events with their grandparents.
The damage inflicted from lockdowns is felt from our bank accounts to our spirits. We are ready for that healthier tomorrow, and the good news is that we can see that day sooner thanks to the groundbreaking development of the COVID-19 vaccine. Through global cooperation of scientists, researchers, and medical experts, an extraordinarily effective and safe vaccine is now available **(5)**.
The speed of development was only possible due to sharing of research on a scale we have never seen before. Scientists had a head start thanks to prior research done on SARS and MERS.

Figure 9: Example 105 from test dataset labeled by GPT4

only the ingredients they need to be safe and effective **(4)**. A note on vaccine safety
Vaccines go through comprehensive safety and effectiveness testing **(4)**. The Food and Drug Administration (FDA) looks at the results of these tests to decide whether to license the vaccine for use in the United States **(4)**.
Learn more about vaccine safety.
Each ingredient in a vaccine serves a specific purpose. For example, vaccine ingredients may **(4)**:
Help provide immunity (protection) against a specific disease **(4)**
Help keep the vaccine safe and long lasting **(4)**
Be used during the production of the vaccine Ingredients provide immunity **(4)**
Vaccines include ingredients to help your immune system respond and build immunity to a specific disease **(4)**.
For example: Antigens are very small amounts of weak or dead germs that can cause diseases.
They help your immune system learn how to fight off infections faster and more effectively.
The flu virus is an example of an antigen.
Adjuvants , which are in some vaccines, are substances that help your immune system respond more strongly to a vaccine. This increases your immunity against the disease.
Aluminum is an example of an adjuvant **(4)**.
Learn more about how vaccines provide immunity.

Figure 10: Example 114 from test dataset labeled through nichesourcing with Reasons, Stances and Scientific Authorities

only the ingredients they need to be safe and effective. A note on vaccine safety Vaccines go through comprehensive safety and effectiveness testing. The Food and Drug Administration (FDA) looks at the results of these tests to decide whether to license the vaccine for use in the United States.
Learn more about vaccine safety.
Each ingredient in a vaccine serves a specific purpose. For example, vaccine ingredients may:
Help provide immunity (protection) against a specific disease
Help keep the vaccine safe and long lasting
Be used during the production of the vaccine Ingredients provide immunity
Vaccines include ingredients to help your immune system respond and build immunity to a specific disease.
For example: Antigens are very small amounts of weak or dead germs that can cause diseases.
They help your immune system learn how to fight off infections faster and more effectively.
The flu virus is an example of an antigen.
Adjuvants , which are in some vaccines, are substances that help your immune system respond more strongly to a vaccine. This increases your immunity against the disease.
Aluminum is an example of an adjuvant.
Learn more about how vaccines provide immunity.

Figure 11: Example 114 from test dataset labeled by our finetuned Longformer for the task of detecting Reasons

only the ingredients they need to be safe and effective **(5)**. A note on vaccine safety Vaccines go through comprehensive safety and effectiveness testing **(4)**. The Food and Drug Administration (FDA) looks at the results of these tests to decide whether to license the vaccine for use in the United States.
Learn more about vaccine safety.
Each ingredient in a vaccine serves a specific purpose. For example, vaccine ingredients may **(4)**:
Help provide immunity (protection) against a specific disease **(5)**
Help keep the vaccine safe and long lasting **(5)**
Be used during the production of the vaccine **(4)** Ingredients provide immunity **(5)**
Vaccines include ingredients to help your immune system respond and build immunity to a specific disease **(5)**.
For example: Antigens are very small amounts of weak or dead germs that can cause diseases **(5)**.
They help your immune system learn how to fight off infections faster and more effectively **(5)**.
The flu virus is an example of an antigen.
Adjuvants , which are in some vaccines, are substances that help your immune system respond more strongly to a vaccine **(4)**. This increases your immunity against the disease **(5)**. Aluminum is an example of an adjuvant.
Learn more about how vaccines provide immunity.

Figure 12: Example 114 from test dataset labeled by our finetuned Longformer for the task of detecting Reasons and their Stance

only the ingredients they need to be safe and effective. A note on vaccine safety Vaccines go through comprehensive safety and effectiveness testing. The Food and Drug Administration (FDA) looks at the results of these tests to decide whether to license the vaccine for use in the United States.
Learn more about vaccine safety.
Each ingredient in a vaccine serves a specific purpose. For example, vaccine ingredients may:
Help provide immunity (protection) against a specific disease
Help keep the vaccine safe and long lasting
Be used during the production of the vaccine Ingredients provide immunity
Vaccines include ingredients to help your immune system respond and build immunity to a specific disease.
For example: Antigens are very small amounts of weak or dead germs that can cause diseases.
They help your immune system learn how to fight off infections faster and more effectively.
The flu virus is an example of an antigen.
Adjuvants , which are in some vaccines, are substances that help your immune system respond more strongly to a vaccine. This increases your immunity against the disease.
Aluminum is an example of an adjuvant.
Learn more about how vaccines provide immunity.

Figure 13: Example 114 from test dataset labeled by our finetuned Roberta-base for the task of detecting Scientific Authorities

Your Price: $10.00 - 100% Satisfaction Guaranteed - 90-Day Money-Back Guarantee - Made in USA
In this 1+ Hour Presentation You'll Learn - Dangerous ingredients in vaccines - Common vaccine injuries - Our 10 step plan to detox from vaccines - How to legally opt out of vaccines in every state - And so much more… **(1)**
Order NOW and Get Instant Access!

Figure 14: Example 160 from test dataset labeled by a human annotator showing a Reason with a Strong stance Against vaccination

spoke to my consultant yesterday about the vaccine , even though we had the virus back in April we did not show any antibodies through the Roche test As for the vaccine , he said it changes every day if we can have it , to do with if its a live vaccine , the pfizer they put a preservatives in that can cause a reactions , similar to what they used to use in some biologics like Enbrel , the preservative called Sulfites0 **(2)**
- Hi, Latest news is that anyone with a history of an allergic reaction or who carry an adrenaline autoinjector should not be given the vaccine, so far no news for those who are on immunosuppresent medication Yvonne x0 - Australia seems to be more upfront **(2)**.
"Because vaccines are tested in healthy populations first, immunocompromised people were also excluded from the Astra-Zeneca and Moderna phase 3 trials, which means safety and efficacy data in these groups is limited **(2)**.
This doesn't mean people who are immunocompromised will be excluded from getting COVID-19 vaccines — it just means health authorities and regulators may wait for further safety data (which is now being collected in clinical trials) before they recommend immunisation to these groups **(4)**.
According to Bruce Thompson, dean of health sciences at Swinburne University, Australia is in a fortunate position in that it has more time to consider the data."We're lucky in Australia that we don't have a lot of the virus, so we can wait a little bit ... and hopefully some of this science catches up," he said **(2)**

Figure 15: Example annotated by human annotator showing a weak stance against vaccination based on scientific debate about immunosuppressed patients

Today, Mayor John Tory announced that the City of Toronto, in partnership with the Governments of Canada and Ontario, has helped bring a new vaccine manufacturing facility to Toronto **(4)**. The new Sanofi facility will be located on their existing site at 1755 Steeles Ave. W. in North York. This new facility will strengthen Canada's domestic vaccine manufacturing capacity and future pandemic preparedness efforts. After design, construction, testing and qualification of the new facility and its equipment, it is expected to be operational by 2026. Through the City's Gold Star program, the municipal government will help Sanofi navigate and expedite the development review and approval process for this important new facility. The new facility may also be eligible for Toronto's Imagination, Manufacturing, Innovation and Technology (IMIT) Financial Incentive Program, which provides grants to support new construction in targeted employment sectors, including biomedical operations and manufacturing. Currently, Sanofi's Toronto Site produces millions of doses annually of life-saving vaccines **(5)** (e.g. against whooping cough, polio, diphtheria and tetanus, among others) to more than 60 countries worldwide - including Canada - and employs 1,500 individuals in industrial affairs, commercial operations and research and development **(5)**. With this new investment, Sanofi will continue to invest in research and development, create a combination of 300 new indirect and direct jobs, and maintain their strong biomanufacturing footprint and highly skilled jobs in Canada. The establishment of a new end-to-end influenza vaccine production facility in Toronto will help to enhance the health of Canadians and secure priority access to domestically manufactured vaccine supply, while providing employment opportunities for residents and adding to the fast-growing health and life sciences sector in the Toronto region **(4)**.

Figure 16: Example 784 from test dataset labeled by a human annotator showing

Today, Mayor John Tory announced that the City of Toronto, in partnership with the Governments of Canada and Ontario, has helped bring a new vaccine manufacturing facility to Toronto. The new Sanofi facility will be located on their existing site at 1755 Steeles Ave. W. in North York. This new facility will strengthen Canada's domestic vaccine manufacturing capacity and future pandemic preparedness efforts. After design, construction, testing and qualification of the new facility and its equipment, it is expected to be operational by 2026. Through the City's Gold Star program, the municipal government will help Sanofi navigate and expedite the development review and approval process for this important new facility. The new facility may also be eligible for Toronto's Imagination, Manufacturing, Innovation and Technology (IMIT) Financial Incentive Program, which provides grants to support new construction in targeted employment sectors, including biomedical operations and manufacturing. Currently, Sanofi's Toronto Site produces millions of doses annually of life-saving vaccines (e.g. against whooping cough, polio, diphtheria and tetanus, among others) to more than 60 countries worldwide - including Canada - and employs 1,500 individuals in industrial affairs, commercial operations and research and development **(5)**. With this new investment, Sanofi will continue to invest in research and development, create a combination of 300 new indirect and direct jobs, and maintain their strong biomanufacturing footprint and highly skilled jobs in Canada. The establishment of a new end-to-end influenza vaccine production facility in Toronto will help to enhance the health of Canadians and secure priority access to domestically manufactured vaccine supply, while providing employment opportunities for residents and adding to the fast-growing health and life sciences sector in the Toronto region.

Figure 17: Example 784 from test dataset labeled by a finetuned longformer for the task of detecting Reasons

Today, Mayor John Tory announced that the City of Toronto, in partnership with the Governments of Canada and Ontario, has helped bring a new vaccine manufacturing facility to Toronto. The new Sanofi facility will be located on their existing site at 1755 Steeles Ave. W. in North York. This new facility will strengthen Canada's domestic vaccine manufacturing capacity and future pandemic preparedness efforts. After design, construction, testing and qualification of the new facility and its equipment, it is expected to be operational by 2026. Through the City's Gold Star program, the municipal government will help Sanofi navigate and expedite the development review and approval process for this important new facility. The new facility may also be eligible for Toronto's Imagination, Manufacturing, Innovation and Technology (IMIT) Financial Incentive Program, which provides grants to support new construction in targeted employment sectors, including biomedical operations and manufacturing. Currently, Sanofi's Toronto Site produces millions of doses annually of life-saving vaccines (e.g. against whooping cough, polio, diphtheria and tetanus, among others) to more than 60 countries worldwide - including Canada - and employs 1,500 individuals in industrial affairs, commercial operations and research and development. With this new investment, Sanofi will continue to invest in research and development, create a combination of 300 new indirect and direct jobs, and maintain their strong biomanufacturing footprint and highly skilled jobs in Canada. The establishment of a new end-to-end influenza vaccine production facility in Toronto will help to enhance the health of Canadians and secure priority access to domestically manufactured vaccine supply, while providing employment opportunities for residents and adding to the fast-growing health and life sciences sector in the Toronto region.

Figure 18: Example 784 from test dataset labeled by a finetuned longformer for the task of detecting Reasons and their Stance

Today, Mayor John Tory announced that the City of Toronto, in partnership with the Governments of Canada and Ontario, has helped bring a new vaccine manufacturing facility to Toronto. The new Sanofi facility will be located on their existing site at 1755 Steeles Ave. W. in North York. This new facility will strengthen Canada's domestic vaccine manufacturing capacity and future pandemic preparedness efforts. After design, construction, testing and qualification of the new facility and its equipment, it is expected to be operational by 2026. Through the City's Gold Star program, the municipal government will help Sanofi navigate and expedite the development review and approval process for this important new facility. The new facility may also be eligible for Toronto's Imagination, Manufacturing, Innovation and Technology (IMIT) Financial Incentive Program, which provides grants to support new construction in targeted employment sectors, including biomedical operations and manufacturing. Currently, Sanofi's Toronto Site produces millions of doses annually of life-saving vaccines (e.g. against whooping cough, polio, diphtheria and tetanus, among others) to more than 60 countries worldwide - including Canada **(3)** - and employs 1,500 individuals in industrial affairs, commercial operations and research and development. With this new investment, Sanofi will continue to invest in research and development, create a combination of 300 new indirect and direct jobs, and maintain their strong biomanufacturing footprint and highly skilled jobs in Canada. The establishment of a new end-to-end influenza vaccine production facility in Toronto will help to enhance the health of Canadians and secure priority access to domestically manufactured vaccine supply, while providing employment opportunities for residents and adding to the fast-growing health and life sciences sector in the Toronto region.

Figure 19: Example 784 from test dataset labeled by a human annotator showing a Reason with a Strong stance Against vaccination

Today, Mayor John Tory announced that the City of Toronto, in partnership with the Governments of Canada and Ontario, has helped bring a new vaccine manufacturing facility to Toronto. The new Sanofi facility will be located on their existing site at 1755 Steeles Ave. W. in North York. This new facility will strengthen Canada's domestic vaccine manufacturing capacity and future pandemic preparedness efforts. After design, construction, testing and qualification of the new facility and its equipment, it is expected to be operational by 2026. Through the City's Gold Star program, the municipal government will help Sanofi navigate and expedite the development review and approval process for this important new facility. The new facility may also be eligible for Toronto's Imagination, Manufacturing, Innovation and Technology (IMIT) Financial Incentive Program, which provides grants to support new construction in targeted employment sectors, including biomedical operations and manufacturing. Currently, Sanofi's Toronto Site produces millions of doses annually of life-saving vaccines (e.g. against whooping cough, polio, diphtheria and tetanus, among others) to more than 60 countries worldwide - including Canada - and employs 1,500 individuals in industrial affairs, commercial operations and research and development. With this new investment, Sanofi will continue to invest in research and development, create a combination of 300 new indirect and direct jobs, and maintain their strong biomanufacturing footprint and highly skilled jobs in Canada. The establishment of a new end-to-end influenza vaccine production facility in Toronto will help to enhance the health of Canadians and secure priority access to domestically manufactured vaccine supply, while providing employment opportunities for residents and adding to the fast-growing health and life sciences sector in the Toronto region.

Figure 20: Example 784 from test dataset labeled by a finetuned Roberta model for the task of detecting Scientific Authorities

Today, Mayor John Tory announced that the City of Toronto, in partnership with the Governments of Canada and Ontario, has helped bring a new vaccine manufacturing facility to Toronto. The new Sanofi facility will be located on their existing site at 1755 Steeles Ave. W. in North York. This new facility will strengthen Canada's domestic vaccine manufacturing capacity and future pandemic preparedness efforts **(4)**. After design, construction, testing and qualification of the new facility and its equipment, it is expected to be operational by 2026. Through the City's Gold Star program, the municipal government will help Sanofi navigate and expedite the development review and approval process for this important new facility. The new facility may also be eligible for Toronto's Imagination, Manufacturing, Innovation and Technology (IMIT) Financial Incentive Program, which provides grants to support new construction in targeted employment sectors, including biomedical operations and manufacturing. Currently, Sanofi's Toronto Site produces millions of doses annually of life-saving vaccines (e.g. against whooping cough, polio, diphtheria and tetanus, among others) to more than 60 countries worldwide - including Canada - and employs 1,500 individuals in industrial affairs, commercial operations and research and development **(5)**. With this new investment, Sanofi will continue to invest in research and development, create a combination of 300 new indirect and direct jobs, and maintain their strong biomanufacturing footprint and highly skilled jobs in Canada. The establishment of a new end-to-end influenza vaccine production facility in Toronto will help to enhance the health of Canadians and secure priority access to domestically manufactured vaccine supply, while providing employment opportunities for residents and adding to the fast-growing health and life sciences sector in the Toronto region **(5)**.

Figure 21: Example 784 from test dataset labeled by a GPT4