Addressing Popular Concerns regarding COVID-19 Vaccination with Natural Language Argumentation Dialogues

Lisa A. CHALAGUINE [a] and Anthony HUNTER [a]
[a] *Department of Computer Science, University College London, London, UK*

Abstract. Chatbots have the potential of being used as dialogical argumentation systems for behaviour change applications. They thereby offer a cost-effective and scalable alternative to in-person consultations with health professionals that users could engage in from the comfort of their own home. During events like the global COVID-19 pandemic, it is even more important than usual that people are well informed and make conscious decisions that benefit themselves. Getting a COVID-19 vaccine is a prime example of a behaviour that benefits the individual, as well as society as a whole. In this paper, we present a chatbot that engages in dialogues with users who do not want to get vaccinated, with the goal to persuade them to change their stance and get a vaccine. The chatbot is equipped with a small repository of arguments that it uses to counter user arguments on why the user is reluctant to get a vaccine. We evaluate our chatbot in a study with participants.

Keywords. chatbots, argumentative persuasion systems, computational persuasion, natural language argumentation, knowledge base construction

1. Introduction

During events like the global COVID-19 pandemic, it is even more important than usual that people are well informed and make conscious decisions that benefit themselves and society. One such example is the willingness to get a vaccine. Vaccines have historically proven to be highly successful and cost-effective public health tools for disease prevention [19]. But the effectiveness of a vaccine in controlling the spread of COVID-19 depends on the willingness to get vaccinated in the general population. A sufficiently high vaccine coverage may generate herd immunity, which will protect everyone, including those particularly susceptible to the virus [11]. However, a barrier to reaching herd immunity is the prevalence of people who refuse or are hesitant to take vaccines [14,20]. For example, the most recent numbers from YouGov surveys on vaccine hesitancy from late March 2021 show that whereas the numbers in the UK are quite high (around 86%), the numbers in, for example, neighbouring France are much lower, at only 49%. In the USA they are a bit higher, at around 59% [1]. Interviewing all those people who refuse to vaccinate in person and trying to convince them to get the vaccine, would be an impossible task.

This problem can be tackled as an argumentation problem: Arguments can be used to provide information and overturn misconceptions [13]. They are an essential part of sensible discussions on controversial and problematic topics. However, despite an increasing body of literature on computational models of argument, there is still a lack of practical applications. Conversational agents, also known as chatbots, have the potential of being used as dialogical argumentation systems for behaviour change applications by applying computational models of argument. A chatbot could engage in an argumentative dialogue with people over the internet from the comfort and safety of their own home, trying to persuade them to get a COVID-19 vaccine. Chatbots could thereby offer a cost-effective and scalable alternative to in-person consultations with health professionals. In order to represent arguments and concerns in the chatbot's knowledge base, we introduce a variant of an argument graph called a *concern-argument graph*.

The problem with controversial topics like the COVID-19 vaccine is that people block out information they disagree with: through creating social media echo chambers, reading partisan news, or only surrounding themselves with like-minded people. A recent paper [16] found that people select more belief-consistent information and perceive belief-confirming information as more credible, useful, and convincing when searching for online health information, also known as *confirmation bias*. Therefore, a dialogue with someone (or something - like a chatbot) could expose people to new information and potentially have a positive effect on their decision-making process.

However, different people worry about different things and hence arguments for not getting a COVID-19 vaccine will vary in a population. One person might be worried about the potential side effects of a newly developed vaccine, whereas someone else might think that he or she does not need a vaccine because they are young and healthy. A chatbot could address those different concerns by providing counterarguments tailored to the different user arguments and during the course of an argumentative dialogue, try to persuade the user to change their stance about getting vaccinated.

Whilst this seems like an obvious idea, there is a lack of a general framework for using concerns in making strategic choices of move in dialogical argumentation systems. A concern is not the same as a value in value-based argumentation frameworks [5,4], since values are used for ignoring attacks by counterarguments where the value of the attacking argument is lower ranked than the attacked one [12]. In our case, however, the chatbot's counterargument addresses the *same* concern as the user's argument, instead of presenting an argument that raises an *opposing* value.

In this paper, we present a chatbot that engages in persuasive dialogues with users who are reluctant to get a COVID-19 vaccine. We show that given a novel domain, like the COVID-19 pandemic and the associated vaccine development, it requires a relatively small repository of counterarguments to address the majority of possible arguments people might have for not getting the vaccine.

The rest of the paper is structured as follows: Section 2 gives some background theory on concerns and the use of argument graphs to construct a chatbot knowledge base; Section 3 gives the aim of the paper and the hypotheses; Section 4 describes the chatbot architecture that was used for the experiments; Section 5 describes the experiments that were conducted with the chatbot, Section 6 presents the results, and in Section 7 we discuss and conclude our findings.

2. Conceptualising Argumentation

2.1. Using Arguments and Concerns for a Persuasive Chatbot

In order for the chatbot to be able to engage in persuasive dialogues, it needs to be equipped with arguments from both perspectives. It needs arguments for (not) engaging in a certain behaviour, that could potentially be given by the user, and arguments that attack the user's arguments (counterarguments). Furthermore, the chatbot should also be able to identify the *concerns* of the user, a concern being a matter of interest or importance to the user. We have shown in a previous study unrelated to healthcare [8], that arguments and counterarguments can be crowdsourced, represented as a graph, and used as the chatbot's knowledge base. Further, we have shown that the chatbot can automatically identify the concerns of the user that she raises in her arguments during the chat, in order for the chatbot to present counterarguments that address the user's concerns.

Taking the user's concerns into account when presenting an argument is important. The chatbot might present a perfectly valid argument that the user might not even disagree with. However, the chatbot's argument may have no impact on her stance, if the argument does not address her concerns. Whereas, if the chatbot presents a counterargument that addresses her concern, the user is more likely to be convinced. In order to choose a suitable counterargument in a given dialogue, our chatbot tries to identify the concern of the user argument and counter with an argument that addresses the user's concern.

To illustrate how concerns arise in argumentation, and how they can be harnessed to improve persuasion, consider a person who is reluctant to the idea of getting a COVID-19 vaccine due to the short time it took to be developed compared to other vaccines. The chatbot (that argues for getting a COVID-19 vaccine) could choose one of the following arguments:

- Option 1: *A vaccine is the only safe way to create herd immunity which is necessary to stop the virus from spreading. This will protect us all from getting the virus, as well as those who, for some reason, cannot get the vaccine.*
- Option 2: *Yes, the vaccine was developed fast. That's because a lot of funding and research priority is currently directed at developing COVID-19 vaccines around the world. As a result, the process is being sped up, allowing for more clinical trials to take place in a shorter period of time.*

Both arguments are perfectly reasonable. However, option 2 addresses the concern regarding the short time frame by acknowledging it and providing an argument on why the user should not worry about it, whereas option 1 does not, even though it provides a valid argument for getting a COVID-19 vaccine.

2.2. Using an Argument Graph as Chatbot Knowledge Base

Argument graphs as proposed by Dung [10] are directed graphs where each node denotes an argument and each arc denotes one argument attacking another. Such graphs provide a useful representation to study attack and support relationships of a given set of arguments.

In our previous studies we have used chatbots to persuade people to cycle more, instead of driving or using public transport [12]; to reduce meat consumption [9], and to change their stance on UK university fees [8]. The arguments were either compiled and edited by hand by the researchers [12] or crowdsourced [9,8] and stored as directed graphs in the chatbot's knowledge base. In [7] we described a method for automatically acquiring a large argument graph via crowdsourcing.

In this study, we used a hybrid approach where we crowdsourced the arguments that people have for not taking a COVID-19 vaccine and hand-crafted the counterarguments for the chatbot ourselves to avoid including invalid or emotionally-loaded arguments in the argument graph. Crowdsourcing arguments provides insight into the reasons for people's behaviour, and an indication of what sort of free natural language input the chatbot needs to handle.

We refrained from crowdsourcing counterarguments because COVID-19 is a serious health issue with global impacts. It is a new domain and research on this disease contains many uncertainties, and scientists still do not know many crucial aspects regarding the virus and its transmission characteristics. Moreover, it is the first time that humans are being injected with vaccines based on mRNA technology. All these form a highly uncertain information landscape and we believe that a carefully curated knowledge base is better than a crowdsourced one in such a critical communication framework.

3. Hypotheses

In this paper, we present a chatbot that utilises a set of arguments for taking a COVID-19 vaccine as a knowledge base. The chatbot uses concerns to make strategic choices of moves in order to engage in argumentative dialogues with users to persuade them to get the vaccine.

Given this setting, we want to address the following three questions: Firstly, whether a small and shallow argument graph is enough to counter the majority of arguments people might have for not getting the vaccine and thereby create persuasive dialogues. Secondly, whether it is possible by only identifying the concern of a user argument to give a suitable counterargument. And finally, whether an interactive chatbot is more persuasive than a static web page that presents arguments for getting the vaccine. We summarise these points in the following three hypotheses:

- **H1** Given a novel domain (i.e. a domain which is relatively new to the user of the chatbot and for which their knowledge and opinions might be limited), a small set of arguments (between 30-50 arguments) can be used to represent most of the possible arguments that a set of normal users would know and appropriate counterarguments, and can be utilised by a chatbot to create persuasive dialogues, meaning that the stance of the user changes after the chat.
- **H2** Given a novel domain, the arguments that address the same concern are sufficiently similar to allow for the provision of suitable counterarguments just by identifying the concern of the arguments.
- **H3** An argumentative dialogue with an interactive chatbot has a higher persuasive effect than presenting the same arguments on a static web page for people to read.

Further, we were also interested in whether during the chats new concerns could be identified which were not raised in the crowdsourced arguments that were used to construct the chatbot's knowledge base.

We would like to note that the authors of this paper are neither psychologists, nor health professionals and that in this work we are (1) not taking any personality traits of the user attributes into account in order to evaluate what sort of argument might be more effective for this particular person, (2) do not compare the persuasiveness of the chatbot's arguments to other potential arguments, and (3) do not incorporate any other methods of persuasion apart from argumentation. Our aim is to present a prototype chatbot that can be used to convince people to get a COVID-19 vaccine using argumentation, and leave the aforementioned issues for future work.

In the remainder of this paper, we describe the design of our chatbot that was used for the argumentative dialogues and explain the experiments conducted with the chatbot in order to test our hypotheses.

4. Chatbot Design

In this section, we describe the acquisition of arguments used to construct the chatbot's knowledge base, and the concern classifier, used by the chatbot to identify the concerns of the incoming user arguments.

4.1. Knowledge Base Construction

To construct the chatbot's knowledge base which consists of a concern-argument graph, which we define below, we recruited 100 participants via Prolific[1], which is an online recruiting platform for scientific research studies, and asked them to provide three arguments against getting a COVID-19 vaccine. We identified 7 concerns that were raised by the majority of the 300 crowdsourced arguments by inspecting the most common, meaningful words, namely: **short-term side effects** of the vaccine, **long-term side effects** of the vaccine, its **fast development**, the **mutation** of the virus, the **safety** of the vaccine, comparison of COVID-19 to the **flu** and downplaying its danger, and **young** people believing they do not need a vaccine.

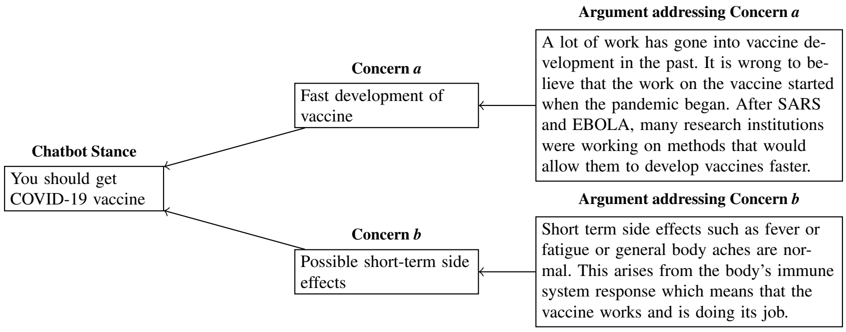
The arguments given by users for each of the identified concerns were quite similar as the following example demonstrates:

- *I will not get a COVID-19 vaccine because of its potential side effects.*
- *It's a new vaccine, so we don't yet know what the side effects are.*
- *The vaccine may have a lot of side effects that could be more dangerous than COVID-19 itself.*

The arguments regarding side effects are very similar, and could all be countered with the same argument, for example, that there is high scrutiny over the research on those particular vaccines and nobody would allow giving it to the public if it was unsafe. Whereas, drafting individual counterarguments would make the graph unnecessarily big and might result in the inclusion of many similar counterarguments. We, therefore, struc-

[1] `https://prolific.co`

Figure 1. Part of the chatbot's knowledge base with concerns *a* and *b* representing clusters of arguments in level 1 (concerned about insufficient testing of the vaccine and possible side effects respectively) attacking the chatbot's goal argument, and arguments in level 2 addressing the concerns.



tured the knowledge base as a *concern-argument graph* which we define as follows:

A **concern-argument graph** is an acyclic graph (N, E, L) where N is a set of nodes, E is a set of edges, and L is a labelling function such that

- N can be partitioned into a set of arguments A and a set of concerns C (i.e. $N = A \cup C$ and $A \cap C = \emptyset$);
- the root of the tree, denoted ρ, is an argument in A;
- for each argument $\alpha \in A \setminus \{\rho\}$, the labelling function assigns a set of concerns to α (i.e. $L(\alpha) \subseteq C$);
- E is the smallest set satisfying the following conditions:
  - $*$ for each concern $\sigma \in C$, there is an edge $(\sigma, \rho) \in E$
  - $*$ for each argument $\alpha \in A \setminus \{\rho\}$, if $\sigma \in L(\alpha)$, there is an edge (α, σ)

So a concern-argument graph can be regarded as a compacted version of a three-level acyclic argument graph (i.e. an argument graph with an argument at the root, counterarguments to the root argument, and counter-counterarguments to the counterarguments). Furthermore, in the concern-argument graph, instead of counterarguments, we use concerns as place-holders for a set of similar arguments raising that concern. This information about concerns we will use in the dialogue strategy.

The resulting graph is, therefore, much smaller and shallower than the traditional argument graph that includes both arguments and counterarguments. Figure 1 shows an example of such a graph, where the goal argument is being attacked by two concerns and each concern is being attacked by one counterargument. One could, of course, extend the concern-argument graph and include more concerns that attack specific counterarguments in level 4, and counterarguments that attack these in level 5. But we aimed to investigate whether a simple 3-layer model would be sufficient in this particular domain where people are more likely to present a new argument, instead of countering the chatbot's counterargument.

The counterarguments for level 3 in the concern-argument graph were carefully researched using reliable sources. Some arguments included links to the NHS website that lists potential short-term side effects from the vaccine[2], or videos that explain how the

[2]https://www.nhs.uk/conditions/coronavirus-covid-19/coronavirus-vaccination/coronavirus-vaccine/

vaccine was developed so quickly without compromising its safety[3]. As already stated above, we did not individually evaluate the persuasiveness of the arguments, as this is out of the scope of this study.

The chatbot also needed default arguments for getting a COVID-19 vaccine that it could use in case the concern of the user argument could not be identified. These arguments are therefore not counterarguments in the traditional sense as they do not directly counter the user argument but instead "change topic" and present a new, important issue in the debate. We also added phrases like *"Ok but", "Have you considered that"* and *"Nevertheless"* to the beginning of the default arguments to indicate that a deviation from the topic occurs. This way the dialogue would resemble argumentation as it would happen between two people: if two human agents engage in an argumentative dialogue, just because one presents an argument the other cannot counter, the dialogue does not necessarily end at that point. The other agent might switch topics and present a new argument he or she believes in, without referencing and directly countering the previous argument. Two out of the three default arguments stated the importance of the vaccine in order to reach *herd immunity*.

Our initial concern-argument graph consisted of 7 concerns each with 2-4 counterarguments, with 19 leaves in total in the concern-argument graph. Additionally, the chatbot was equipped with three default arguments it could use in case no concern could be identified.

4.2. Understanding the user input

The initial move is the root argument presented by the chatbot, then the user gives a counterargument which is analysed to determine the concern. The chatbot identified the concern of the user argument using a multinomial logistic regression and a binary feature representation of the arguments (one-hot encoded vectors)[4]. For the initial classifier, the crowdsourced arguments were used for training. If the prediction was over 40% in confidence, the argument was labelled with the identified concern which was used by the chatbot to pick a leaf argument from the concern-argument graph that is a counterargument to that concern. If no concern cannot be identified, or all the leaves of the identified concern had been used, a default argument was presented to the user.

5. Experiments

The purpose of the chatbot was to test all three of our hypotheses. Prior to recruiting participants for the study, we ran a survey where we asked people to choose from a scale of 1-5 whether they would get a COVID-19 vaccine. The options were *very unlikely, somewhat unlikely, neither likely nor unlikely, somewhat likely* and *very likely*. We recruited 300 participants from those that chose *very unlikely, somewhat unlikely* and *neither likely nor unlikely*, i.e. those that had a negative or neutral stance. 240 participants chatted with the chatbot and 60 participants were presented a static web page.

[3]All arguments were fact-checked by a medical doctor trainee and a machine learning consultant who works in the pharmaceutical industry

[4]Due to the small amount of data and the use of one-hot encoded vectors, there was no considerable difference when evaluating different classifiers. More sophisticated methods, like pre-trained language models, could not be used due to the lack of data that could be used for pretraining on the given topic.

Before the chat, the users were directed to a Microsoft Form and asked again how likely they would get a COVID-19 vaccine. After submitting their answer they were redirected to a web page where they could begin the chat. The chatbot was composed of a front-end we coded in Javascript and a Python back-end using the Flask web server library. The chatbot started the chat by instructing the user that they could end the chat anytime by sending the word *"quit"* and then asking why the user would not get a COVID-19 vaccine, once it became available to him/her. The user then presented his/her first argument. The chatbot replied with either a counterargument from the concern-argument graph or a default argument, depending on whether it could identify the concern of the user argument. The counterarguments were stored in a Python dictionary with the concerns as the keys and the list of counterarguments that addressed that concern as the values. If the concern could be identified, the chatbot replied with the first counterargument in the list. If the message of the user was less than 7 words in length and contained a negation, the chatbot queried *Why?* or *Why not?* to force the user to expand. This process was repeated with each subsequent argument given by the user. The chatbot would end the chat as soon as all default arguments were used up and no concern could be identified, or all counterarguments that addressed the concern were also used up. At the end of the chat the chatbot presented the user with a link that redirected them to another Microsoft Form where they were asked a series of questions:

1. Did you feel understood by the chatbot? (Yes/No/Sometimes)
2. Did you feel that the chatbot's arguments were relevant? (Yes/No/Some of them)
3. Do you feel like all your concerns were addressed? (The majority of them/None of them/Some of them)
4. How likely would you get a COVID-19 vaccine, once one becomes available to you? (Very unlikely - very likely)

Questions 1-3 were used to test our second hypothesis and judge the relevance, length and quality of the chats, and question 4 was to test our first and third hypotheses and compare the stances of the participants before and after the chat with the chatbot in order to judge persuasiveness. In order to test our third hypothesis, 60 out of the 300[5] participants did not chat with the chatbot but instead were presented the chatbot's 10 most commonly used counterarguments in *persuasive* chats on a static web page.

In order to test whether new concerns can be identified during the chats that were not identified in the crowdsourced arguments, we analysed the chats after every batch of 60 participants (we also recruited participants in batches of 60). By inspecting common, meaningful words we could identify new concerns after each batch of 60 participants and re-train the classifier with enough examples of the new concern, and add suitable counterarguments to the concern-argument graph. We only added a new concern to the chatbot's concern-argument graph if we could automatically identify at least 10 arguments that addressed that concern. The training set was also updated with more training examples to an already existing concern. For example, in the crowdsourced arguments, many people used the word *mutation* whereas in the chats the word *strain* was prevalent.

This way 8 additional concerns could be identified: **death**, that the vaccine does not prevent you from **getting** and **spreading** COVID-19, people claiming they already

[5] Due to limited funding we did not want to split the participants in half but rather collect more data in form of chatlogs, since a web page does not provide data for further research.

Table 1. Breakdown of the 240 participants' stance for getting a COVID-19 vaccine before and after chatting with the chatbot.

	Very unlikely	Somewhat unlikely	Neither likely nor unlikely	Somewhat likely	Very likely
Before	30%	42%	28%	0%	0%
After	22.5%	37.5%	29%	9.5%	1.5%

Table 2. Percentage of the 240 participants who changed their stance after chatting with the chatbot.

Negative to Neutral	Neutral to Positive	Negative to Positive	Total
9%	7.5%	3.5%	20%

had COVID-19, that COVID-19 has a too high of a **survival rate** to be worried about it, that the vaccine may impact **fertility**, that the vaccine might not be **effective**, that the **ingredients** of the vaccine are unknown, and that herd immunity can be created **naturally** by catching the virus and hence no vaccine is needed.

6. Evaluation of the Chatbot

Table 1 shows the stance of the 240 participants who chatted with the chatbot, before and after chatting with the chatbot. We divided the change in stance into 3 categories: a change from negative to neutral (from *very unlikely/ somewhat unlikely* to *neither likely nor unlikely*); a change from neutral to positive (from *neither likely nor unlikely* to *somewhat likely/very likely*); and a change from negative to positive (from *very unlikely/somewhat unlikely* to *somewhat likely/very likely*). We do not consider a change from *very unlikely* to *somewhat unlikely*. Table 2 shows the percentage of the 240 participants who changed their stance by engaging in an argumentative dialogue with the chatbot. 20% of the participants (48 out of 240) had a positive change in stance. This verifies our first hypothesis - that a small, shallow concern-argument graph can be utilised by a chatbot to create persuasive dialogues.

Given that the chatbot did not use natural language generation and was not able to address novel arguments or expand on existing ones by giving more information, and solely relied on correct concern classification, the given results are promising. The length of the chats were on average 12 alternating turns. This means that the chatbot, on average, gave 6 arguments, 3 of which were default arguments and 3 from the graph. Table 3 shows the results for the first three questions. 35% of the participants felt understood by the chatbot and further 41% felt sometimes understood. 32% perceived the chatbot's arguments as relevant and further 55% perceived them as sometimes relevant. 23% felt that the majority of their concerns were addressed and further 54% felt that some of their concerns were addressed. This supports our second hypothesis that only by identifying the concern of an argument, suitable counterarguments can be presented and that the resulting argumentation dialogues are of satisfactory length and quality. An example of

Table 3. Answers to the first three questions by the 240 participants who chatted with the chatbot.

Felt understood			Relevance			Concern addressed		
YES	SOMETIMES	NO	YES	SOME	NO	MAJORITY	SOME	NONE
35%	41%	24%	32%	55%	13%	23%	54%	23%

Table 4. Breakdown of stance for getting a COVID-19 vaccine of the group of 60 participants before chatting with the chatbot, and the group of 60 participants who was presented with a static web page.

	Very unlikely	Somewhat unlikely	Neither likely nor unlikely
Chatbot	60%	33%	7%
Web page	52%	28%	20%

Table 5. Change of stance for the group of 60 participants who chatted with the chatbot, and the group of 60 participants who was presented with a static web page.

	Negative to Neutral	Neutral to Positive	Negative to Positive	Total (no of participants)
Chatbot	12%	5%	2%	18% (11)
Web Page	5%	0%	0%	5% (3)

a chat can be seen in Figure 2. All chatlogs, the data for the concern-argument graph, the concerns and their descriptions, and the code for the chatbot can be found on github [2].

As mentioned in the previous section, we recruited the participants in batches of 60. To evaluate our third hypothesis, the chatbot's persuasive effect compared to a static web page, we compared the results of the 60 participants who were presented with the static web page with a batch of 60 participants who chatted with the chatbot with similar starting distributions of their stance. The fourth (and last) batch of participants had a similar distribution as the batch that was recruited to read the arguments on a static web page. The starting distributions are shown in Table 4 and the change of stance for both groups are shown in Table 5.

We used a Chi-Square test to compare the number of participants who changed their stance after chatting with the chatbot with the number of participants who changed their stance after reading the 10 most common arguments used by the chatbot (this means that the participants who only saw the static web page, on average saw 4 more arguments than those who chatted with the chatbot). The results were statistically significant with a p-value of .023 at $p < .05$. The results, therefore, support our hypothesis that an interactive chatbot is more persuasive than a static web page.

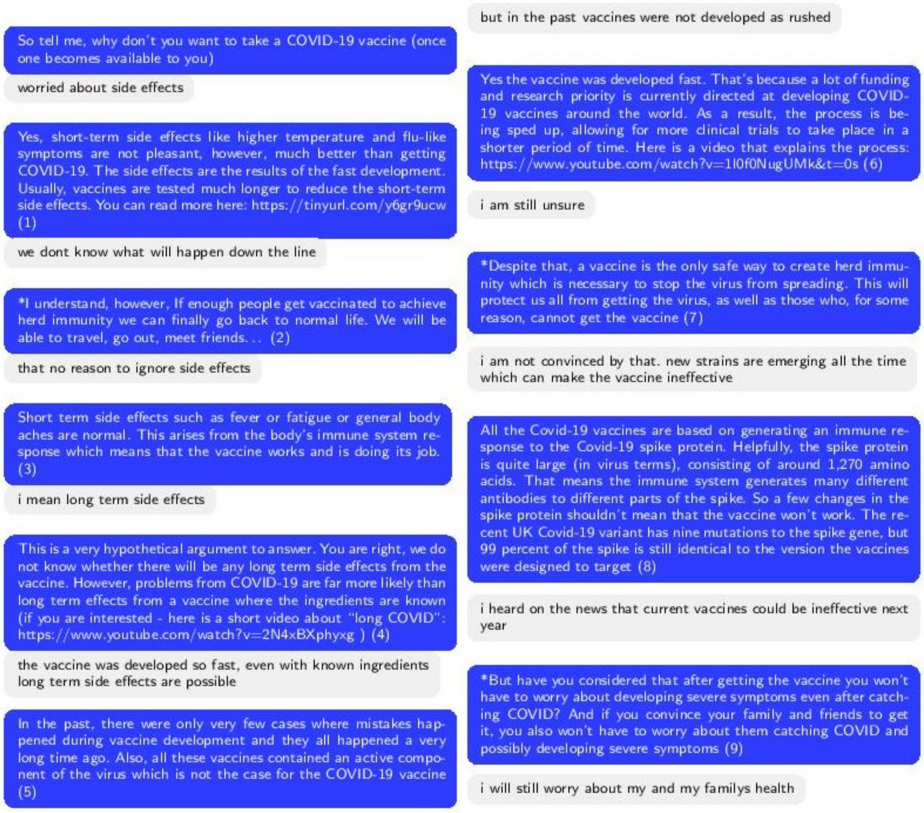
The newly identified concerns mentioned in the previous section were also raised by some arguments that were collected during the initial argument collection described in Section 4.1. So claiming that new concerns could be identified during the chats would be incorrect. Given a large enough sample of crowdsourced arguments, more concerns could have been identified and included (together with appropriate counterarguments). in the concern-argument graph that was used as the chatbot's knowledge base.

Side effects (short and long term) are by far the most popular concern, with 45% of user arguments given during the chats (where a concern could be identified) raising it. This is coherent with previous studies which analysed vaccine hesitancy in France [21], the US [15] and the EU [17]. The second most prevalent concern was about the safety of the vaccine in general (28%), and in the third place, the vaccines fast development and young people believing they do not need one (both 10%).

7. Discussion and Conclusion

The aim of this paper was to present a prototype chatbot that can engage in persuasive dialogues with people who are opposed to the COVID-19 vaccine using computational

Figure 2. Example chat between a participant from the fourth batch. Chatbot arguments are in the dark boxes and user arguments in the light boxes. Default arguments are indicated with an *. The chat begins in the left column and continues in the right one. The participant indicated that he or she was *somewhat unlikely* to get vaccinated before the chat, but changes his/her stance to *neither likely nor unlikely* after the chat.



models of argument. Our contribution in this paper is threefold. Firstly, we have shown that for a new domain, where there exists a lot of uncertainty, a small argument graph can be used to represent most of the possible arguments in this domain. This argument graph can be utilised by a chatbot to create persuasive dialogues, and we presented a method how to acquire and structure such a graph in form of a concern-argument graph. In our previous work [8] the chatbot's knowledge base consisted of an argument graph that included both arguments and counterarguments. The chatbot matched the incoming user argument with a similar argument in the graph (target argument) using cosine similarity of the vector representations of the two arguments (the vectors were created using GloVe word embeddings [18]). The graph was therefore much bigger (containing over 1200 arguments) than the one presented in this paper.

Secondly, we have demonstrated that no sophisticated natural language understanding of the user arguments is needed in order to provide suitable counterarguments that address the majority of the concerns of the users. And thirdly, we have shown that an interactive chatbot has a higher persuasive effect than a static web page.

Further, we have shown that in this domain a concern-argument graph (a three-level acyclic graph) where after the initial move, the chatbot only picks a leaf at every turn or uses a default argument, is enough to generate persuasive dialogues. Using a modest concern-argument graph, as described in this paper, and to not constraint the chatbot to use a larger argument graph that may involve long paths, has two main benefits: firstly, the graph can be constructed with less data than for a larger argument graph; secondly, this allows the chatbot to counter user arguments that are not direct counterarguments to the previously given chatbot argument. This is important, as during the chats people often ignored the chatbot's counterargument but instead gave a new argument on why not to take the vaccine.

It would, of course, be desirable to use more sophisticated natural language processing methods to process the user's input. For that, however, much more data is needed which currently is not available. The chatbot, hence, can only reply to well-phrased arguments that raise common concerns that it can identify. These sort of replies contributed to only 50% of the users' replies. Other types of responses included novel arguments, statements like *"I don't care* or *"I will take my chances"*, emotional accusations (about the government not caring or the chatbot being stupid) and questions about the vaccine. In future work, argumentation will, therefore, only be one component of the chatbot, paired with the ability to answer questions and provide information about COVID-19 vaccines and their associated risks, and a conversational component that can address user statements and emotional responses.

We also want to experiment with different argumentation frameworks and dialogue strategies. A reasonable extension to the current framework would be *bipolar argumentation* [6,3]. In Figure 2, the fourth argument that the chatbot presents addresses the concern of long term side effects. The user replies with an argument that again raises long term side effects. Hence, the fifth argument by the chatbot also addresses that concern. Argument 5 can, therefore, be seen as a supporting argument to argument 4. A potential dialogue strategy for the chatbot could be to use arguments that support the previously given argument by the chatbot, if a concern cannot be identified, instead of giving default arguments. For example, after the user said that he is still unsure, the chatbot could have provided a supporting argument to argument 6 and present the user with another argument that addresses the fast development of the vaccine.

To conclude this paper, we want to emphasise the advantage of using a chatbot for such a task: a chatbot is able to address millions of people at the same time in the comfort of their own home and collect a vast amount of data in a very short time. Our method of analysing the incoming user arguments scales easily and allows obtaining many arguments from different people. The more data comes in, the easier it gets to identify patterns, discover new concerns, and acquire arguments that address these concerns and update the chatbot's concern-argument graph accordingly. This allows us to identify common misconceptions, address the lack of information, and potentially even fake news.

Acknowledgements

This study was partially supported by Noah Castelo from the Alberta School of Business.

References

[1] `https://yougov.co.uk/topics/international/articles-reports/2021/01/12/covid-19-willingness-be-vaccinated`.

[2] `https://github.com/lisanka93/covid-19_vacc_bot`.

[3] L. Amgoud, C. Cayrol, M.-C. Lagasquie-Schiex, and P. Livet. On bipolarity in argumentation frameworks. *International Journal of Intelligent Systems*, 23(10):1062–1093, 2008.

[4] T. Bench-Capon and K. Atkinson. Abstract argumentation and values. In *Argumentation in artificial intelligence*, pages 45–64. Springer, 2009.

[5] T. J. Bench-Capon. Persuasion in practical argument using value-based argumentation frameworks. *Journal of Logic and Computation*, 13(3):429–448, 2003.

[6] C. Cayrol and M.-C. Lagasquie-Schiex. On the acceptability of arguments in bipolar argumentation frameworks. In *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, pages 378–389. Springer, 2005.

[7] L. A. Chalaguine and A. Hunter. Knowledge acquisition and corpus for argumentation-based chatbots. In *Proc. of the 3rd Workshop on Advances In Argumentation In Artificial Intelligence*, pages 1–14. CEUR, 2019.

[8] L. A. Chalaguine and A. Hunter. A persuasive chatbot using a crowd-sourced argument graph and concerns. In *Proc of Computational Models of Argument (COMMA)*, pages 9–20. IOS Press, 2020.

[9] L. A. Chalaguine, A. Hunter, F. L. Hamilton, and H. W. W. Potts. Impact of argument type and concerns in argumentation with a chatbot. In *Proc. of the 31st International Conference on Tools with Artificial Intelligence*, pages 1557–1562. IEEE, 2019.

[10] P. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77(2):321–357, 1995.

[11] P. Fines, K. Eames, and D. Heymann. "Herd immunity": a rough guide. *Clinical Infectious Diseases*, 52(7):911–916, 2011.

[12] E. Hadoux and A. Hunter. Comfort or safety? Gathering and using the concerns of a participant for better persuasion. *Argument & Computation*, 10(2):113–147, 2019.

[13] A. Hunter. Towards a framework for computational persuasion with applications in behaviour change. *Argument and Computation*, 1:15–40, 2018.

[14] N. MacDonald. Vaccine hesitancy: Definition, scope and determinants. *Vaccine*, 33(34):4161–4164, 2015.

[15] J. E. Marco-Franco, N. Guadalajara-Olmeda, S. G. de Julián, and D. Vivas-Consuelo. Covid-19 healthcare planning: Predicting mortality and the role of the herd immunity barrier in the general population. *Sustainability*, 12(13):5228, 2020.

[16] C. S. Meppelink, E. G. Smit, M. L. Fransen, and N. Diviani. "i was right about vaccination": Confirmation bias and health literacy in online health information seeking. *Journal of health communication*, 24(2):129–140, 2019.

[17] S. Neumann-Böhme, N. V. NE, I. Sabat, P. Barros, W. Brouwer, J. van Exel, J. Schreyögg, and T. Stargardt. Once we have it, will we use it? a european survey on willingness to be vaccinated against covid-19. *The European Journal of Health Economics*, 21:977–982, 2020.

[18] J. Pennington, R. Socher, and C. D. Manning. Glove: Global vectors for word representation. In *Empirical Methods in Natural Language Processing (EMNLP)*, page 1532–1543. ACL, 2014.

[19] V. Rémy, N. Largeron, S. Quilici, and S. Carroll. The economic value of vaccination: why prevention is wealth. *Journal of Market Access & Health Policy*, 3(1), 2015.

[20] S. M. Sherman, L. E. Smith, J. Sim, R. Amlot, M. Cutts, H. Dash, G. J. Rubin, and N. Sevdalis. Covid-19 vaccination intention in the uk: Results from the covid-19 vaccination acceptability study (covaccs), a nationally representative cross-sectional survey. *medRxiv*, 2020.

[21] J. K. Ward, C. Alleaume, and P. Peretti-Watel. The french public's attitudes to a future covid-19 vaccine: the politicization of a public health issue. *https://doi.org/10.31235/osf.io/xphe9*, 2020.