Towards an ontology-based medication conversational agent for PrEP and PEP

Muhammad (Tuan) Amith, Licong Cui, Kirk Roberts, Cui Tao
School of Biomedical Informatics
The University of Texas Health Science Center at Houston
Houston, Texas

Abstract

HIV (human immunodeficiency virus) can damage a human's immune system and cause Acquired Immunodeficiency Syndrome (AIDS) which could lead to severe outcomes, including death. While HIV infections have decreased over the last decade, there is still a significant population where the infection permeates. PrEP and PEP are two proven preventive measures introduced that involve periodic dosage to stop the onset of HIV infection. However, the adherence rates for this medication is low in part due to the lack of information about the medication. There exist several communication barriers that prevent patient-provider communication from happening. In this work, we present our ontology-based method for automating the communication of this medication that can be deployed for live conversational agents for PrEP and PEP. This method facilitates a model of automated conversation between the machine and user can also answer relevant questions.

1 Introduction

HIV can cause a dangerous infection that can lead to AIDS, a disease that can lead to severe immunological symptoms and eventual death. Common modes of infection include sexual contact, blood transfusion, or the sharing of drug paraphernalia. While the rates have dropped over the last few decades, HIV infection is not uncommon. For example, there is an infection rate of 2 million globally (World Health Organization, 2017) and 39,782 within the United States (Hess et al., 2018). In addition, a segment of the American population with HIV are unaware of the HIV status (Centers for Disease Control and Prevention, 2016), and therefore at risk of spreading the disease to other individuals.

Advances in medication introduced PrEP and PEP. PEP refers to the use of antiretroviral drugs for people who are HIV-negative after a single high-risk exposure to stop HIV infection, while PrEP is a prevention method for people who are HIV-negative and have high risks of HIV infection. Both of these treatments require consistent adherence to the dosage in order to be fully effective, but adherence is an issue for patients subscribed to it. Providers in particular are concerned about the consistent adherence to PrEP (Wood et al., 2018; Blackstock et al., 2017; Clement et al., 2018).

It has been reported that if PrEP adherence is high, rates of HIV infection will be sizeably reduced (Smith et al., 2015). However, adherence to PrEP is no different than other challenges with medications, such as the patient comprehending the administration of the medication and remembering to take it (American Medical Association, 2016). On top of that, the Centers for Disease and Control (CDC) specifically prescribes periodic counseling, and coordinating with patients on a one-on-one basis (Centers for Disease Control and Prevention, 2014). But time burdens and manpower to conduct counseling pose another challenge (Krakower et al., 2014).

In a previous study, Amith et al. (2019a) utilized an ontology-based method to model the dialogue for the counseling of the HPV vaccine. In this study, we tailor the method for PrEP and PEP counseling with the intent that this could be employed in portable tools for drug users to use. A benefit of using an ontology approach, other than exploiting network-based model for dialogue, is the potential to link the ontology to representations of health behavior models (like the transtheoretical model). Systems that leverage health behavior models, according to Kennedy et al. (2012), have demonstrated to be more impactful on affecting health behaviors of users. Also, an ontology that models dialogue can yield standardization and sharing. Amith et al. (2019b) noted from their

Proceedings of the 1st Workshop on NLP for Medical Conversations, pages 31–40

https://doi.org/10.18653/v1/P17

literature review on PubMed that there is limited ontology-centric studies for health-based dialogue management. Amith et al. (2020) simulation studies have also shown evidence that automated counseling, specifically conversational agents for vaccines, could impact the health attitudes and beliefs that can lead to improved uptake with perceived high usability.

Ontologies are artifacts that represent and encode domain knowledge for machines to understand a domain and their physical environment. According to one school of thought, if machines have symbolic understanding of their domain and environment, it could potentially provide near-autonomous operation of tasks. Imbuing software with autonomous task of dialogue interaction requires some measure of intelligence. Intelligent agents are defined as having *reactive*, *proactive* and *social ability* features (Wooldridge and Jennings, 1995). *Reactive* refers to the software ability for timely response to the environment. *Proactive* refers to the initiative driven aspect of the software to accomplish tasks, and *social ability* involves the software handling external interaction with the environment (virtual or physical). How these qualities manifest vary by the architectural approach (*reactive agents*, *reasoning agents*, etc.) which is beyond the scope of discussion.

Researchers mention the use of internal data models within the architecture of the agents (Wooldridge, 2009). The models' role in the system is to provide the agent with decision making capabilities to perform autonomously in the environment. This would include 1) representing the domain knowledge for the agent, 2) providing information of the surrounding environment of the agent, and 3) cataloging the previous actions of the agent (e.g., for the agent to learn). According to Hadzic and colleagues, these models could be manifested as a group of ontologies (Hadzic et al., 2009). Furthermore, they state some inherit benefits such as producing shared communication models between agents and systems, information retrieval, organization of the agent's task, and analytical and reasoning of the knowledge (Hadzic et al., 2009).

The ontology-based solution also attempts to solve some of the issues with reasoning agents like the *transduction problem* and the *representation/reasoning problem* (Wooldridge, 2009). The *Transduction problem* is how to translate the world or domain that the agent is embodied in into symbolic representations. The *representation and reasoning problem* pertains to the challenge of manipulating the symbolic representations and applying reasoning for the agent. With ontologies, we can model a domain space or the environment using predicate logic that is syntactically encoded into a machine-readable artifact. Within the context of this work, this method maps utterances of the user and the machine to concepts represented in our ontological model. Also, with the availability of reasoners, like Pellet (Sirin et al., 2007) or HermiT (Glimm et al., 2014; Shearer et al., 2008), we can perform reasoning based on the encoded model to generate inferred dialogue context information.

From a natural language processing (NLP) standpoint, dialogue is essentially a sequence of utterances between multiple agents. Our work utilizes a finite state transition network to model the dialogue (Allen, 1995; Jurafsky and Martin, 2000), and then encodes this sequence model of the utterances within the ontology. We also employ some lightweight NLP methods to help the agent discern participant utterances, alongside with the reasoning capacities of the agent. For the design of the dialogue system, we utilize a deterministic and planned approach to automate the counseling versus a generative approach in order to cover certain main points to communicate to the drug user. This gives us the control needed to ensure the conversational agent delivers the appropriate counseling. The dialogue will center around a closed world domain – specific to only PrEP and PEP, and HIV infection. The following sections will cover the development of the conversational agent and discussion through results of a Trindi Tick assessment for dialogue system evaluation and future steps with our work.

2 Methods

2.1 Ontology Models

We developed a series of ontologies to provide the software agent with interaction abilities – to model patient-level information and the dialogue flow for the agent to coordinate the interaction with the user.

Ontology of PrEP and PEP (OPP) For the PrEP and PEP information source we created the Ontology of PrEP and PEP (OPP), using patient-level sources (brochures and websites). The OPP describes basic dosing, benefit and harms, cost, po-

tential users, and other pertinent information that patients would like to know. This ontology provides a knowledge base for atomic facts for the dialogue flow ontology, PHIDO. This early version of OPP has 152 classes, 57 object properties, 23 data properties, and 10 instance individuals.

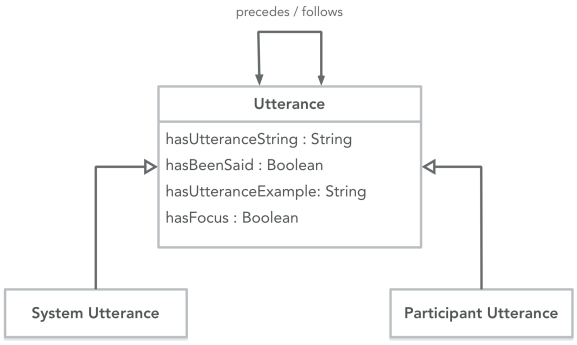

Figure 1: UML diagram of the Utterance class in PHIDO. System and Participant are subclasses of Utterance.

Patient Health Information Ontology Dialogue (PHIDO) The Patient Health Information Ontology Dialogue is an ontology developed in the previous study to model a chain of utterances between the machine (utterances of the system) and the user speaking directly with the machine (utterance of the participant). Figure 1 displays the Utterance class in PHIDO. The parent Utterance class has several data properties that are used to help facilitate the flow the machine's conversation and are linked together using "precedes" or "follows" to indicate precedence of the utterances. PHIDO's TBox level metrics contain 86 classes, 9 object properties, and 5 data properties. Details of the ontology is discussed in the authors' previous study (Amith et al., 2019b).

Essentially, each triple (i.e. predicate) from OPP is utilized by PHIDO to communicate statements about PrEP or PEP (Figure 2). Within PHIDO, an utterance data (instance) is linked to each predicate for the machine to either speak or to help discern utterances spoken by the user.

Figure 3 shows the meta-level description of the dialogue that starts with basic introduction and acclimation of the user with the machine and closing out the counseling. The core goals of the dialogue is to communicate facts (Health Information) and to handle questions at any time for the user (Question Answering). The flow of communication for health information is facilitated by a sub-goal we call Discuss Health Topic (DHT) which is modeled in the PHIDO and allows for population of utterance data that aligns with the concepts in DHT.

2.2 Dialogue System

From a previous study we developed a software engine that uses the aforementioned ontology models (Amith et al., 2019a). The software engine also supports question answering if the ontology model infers a question based on the context of the dialogue flow (e.g., an answer follows a question, a question precedes a point of confusion in the conversation, etc.).

On Figure 4, we demonstrate the execution of the engine using the PHIDO to model the communication of patient-level information to the user. Here the system evokes each patient-level health fact (Health Information) about PrEP, sourced from the OPP. Any question recognized by the engine will attempt to answer the question using an ontology-based question answering (QA) sub-system of the engine. Once the utterance of the user is identified as a question, the type of question, the nouns, and verb phrases are analyzed against the OPP predicates. After selection, ranking, and filtering, the answer is evoked by the system. Details about the implementation is described in (Amith et al., 2019a), and Figure 5 shows an example demonstration with a question and a response from the QA system.

2.2.1 NLP-based Slave Functions

Aside from the finite-state diagram approach for the engine, we also employed some NLP methods as slave functions for the engine to operate - discerning the type of participant utterance for the main dialogue system and comparing the question data with ontology triples for the question-answering subsystem.

Discerning Participant Utterances When capturing input from the user, the engine will need to distinguish the type of user utterance to direct the flow of the dialogue. Each utterance concept was annotated to a set of examples that were representative of its type. For example, the Utterance class of Question has string examples of "could you tell me", "how many", "list", and the Disconfirmation class has examples such as "negative", "never", "none", etc. Using the inputted utterance of the user, the text is compared to the string examples using Monge-Elkan (Monge et al., 1996;

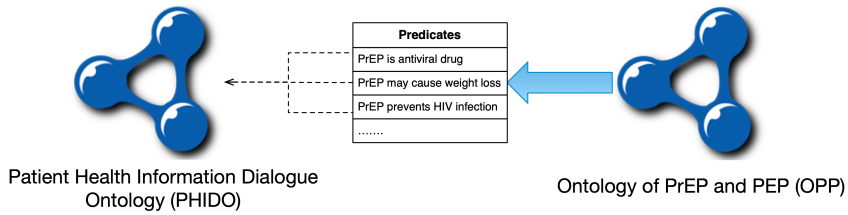

Figure 2: Triples (Predicates) are extracted from the Ontology of PrEP and PEP and used by PHIDO.

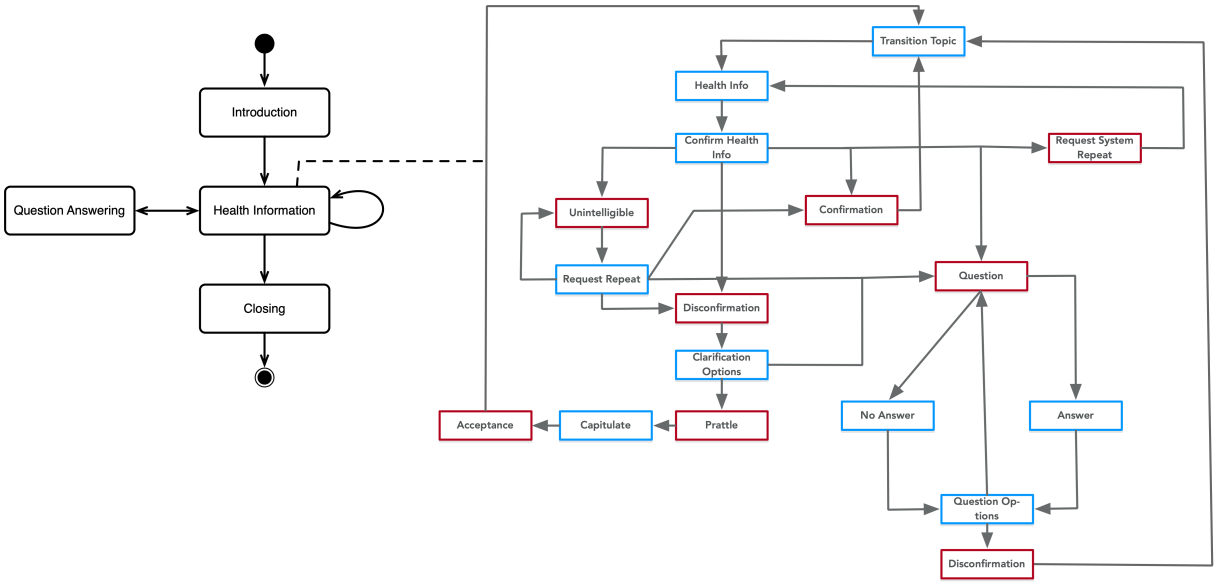

Figure 3: Dialogue flow plan for basic counseling for PrEP and PEP information.

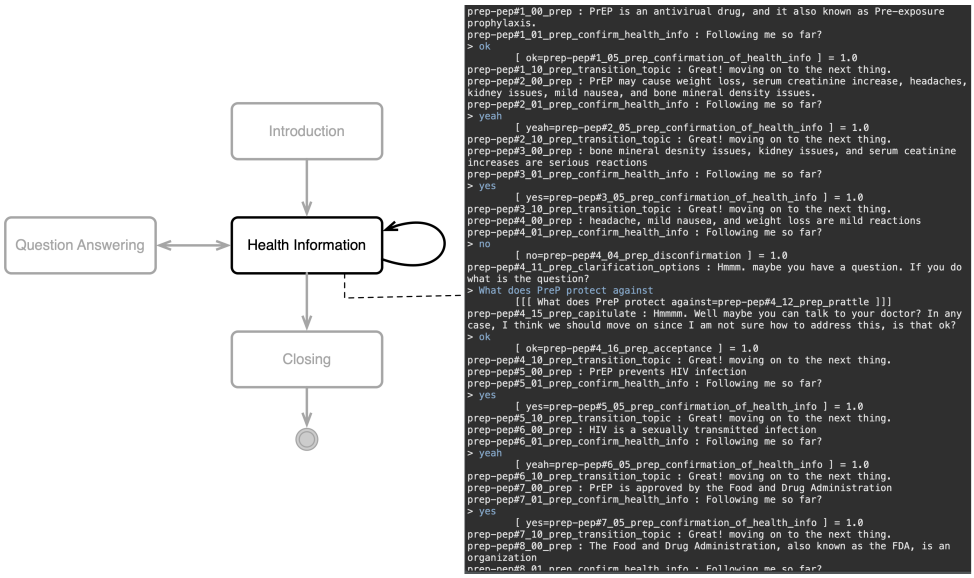

Figure 4: Demonstration of the dialogue engine communicating PrEP information to the user.

Monge and Elkan, 1997) (default implementation from Korstanje (2019)), and a default threshold of 0.85. Matches that do not meet the threshold will fallback to exact string matches based on the beginning of the string.

Definition 2.1 (Participant utterance & examples)**.** *Every participant utterance PU expected by the system contains example data EU_n. EU_n has a number of string text TT_n that are a set of tokens*

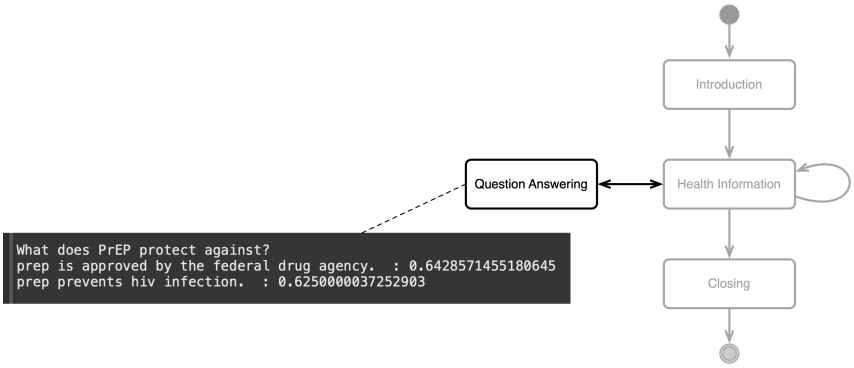

Figure 5: Demonstration of the a PrEP-related question asked with corresponding answers provided. The ranking scores are provided for supplement.

t_n.

$$\forall PU^n \ni EU_n^n$$

$$EU_n^n \in \begin{cases} TT_n, where \\ TT = \{t_1 t_2 \cdots t_n\} \\ t \text{ is a string token} \end{cases}$$

Definition 2.2 (User utterance)**.** *User utterance UU for the dialogue system is a set of string tokens t_n*

$$\forall UU = \{t_1 t_2 \cdots t_n\}$$

Definition 2.3 (Identifying the participant utterance)**.** *To find the exactly identified PU^n the dialogue system attains a comparison $s(x)$ that is the maximum and the greater than a defined threshold TH among all of the example utterances EU^n within each expected utterance PU^n.*

$$PU^n \Rightarrow T_{score} = max(s(x)) > TH$$

$$s(x) = \begin{cases} EU_n^n \cap UU, \\ where \\ EU_n^n = t_1^e t_2^e \cdots t_n^e \end{cases}$$

Defintion 2.3.1 (Identifying the participant utterance)**.** *Assuming that Definition 2.3 fails to find the expected participant utterance PU^n, the dialogue system resorts to finding exact match of the beginning string tokens $EU^{'}$ and $UU^{'}$ from example utterances EU_n^n and the user utterance UU.*

$$PU^n \Rightarrow T_{score} = b(x)$$

$$b(x) = \{EU_n^n \approx UU \Rightarrow EU_n^{'} = UU^{'}\},$$

$$where$$

$$EU_n^{'} \in EU_n^n = \begin{cases} EU_n^{'} = \{t_1^e t_2^e \cdots t_{n-m}^e\} \\ EU_n^n = \{t_1^e t_2^e \cdots t_n^e\} \end{cases}$$

$$UU^{'} \in UU = \begin{cases} UU^{'} = \{t_1 t_2 \cdots t_{n-m}\} \\ UU = \{t_1 t_2 \cdots t_n\} \end{cases}$$

Comparing Question Data and Ontology Triples For the question answering subsystem, the system utilized off the shelf NLP tools like Stanford Core (Manning et al., 2014) to extract data from the question. To preform the matching described in (Amith et al., 2019a), we also utilized a combination of either word embedding using Numberbatch vector model (Speer and Lowry-Duda, 2017) (with Semantic Vectors (Widdows and Cohen, 2010) as the interface layer) or the string similarity methods discussed earlier, and extJWNL (Autayeu, 2016) where we assign a score to each triple from the knowledge base (OPP). For brevity, we applied various rules and thresholds to select and filter triples to present an answer.

Definition 2.4 (Primary Question Data)**.** *Given a question Q, there are a subset of elements NP_n and VP_n (noun phrases and verb phrases) that are essential data $Q^{'}$ for the subsystem.*

$$Q^{'} \in \{NP_n, VP_n\}$$

Definition 2.5 (Triple Assertion)**.** *Within an ontology O, there are assertion triples (ABox) that are composed of elements of subject s_n, predicate p_n, and object o_n to form an assertion triple spo_n.*

$$spo = \{s, p, o\}$$

Definition 2.6 (Essential Ontology Triples)**.** *Given an target ontology O, there are a subset of triple assertion types spo^o, spo^d, spo^c (object property assertions, data property assertions, and class assertions) that are needed O' for the subsystem.*

$$O^{'} \in \{spo_n^o, spo_n^d, spo_n^c\}$$

Definition 2.7 (Assign Score From Comparison)**.** *A similarity score TS is assigned from comparing*

the similarity of question data $D^{'}$ with a triple assertion $spo_n^{\{o,d,c\}}$ from O'. TS is derived from the mean of computing similarity between NP_n with $s_n^{\{o,d,c\}}$ and $o_n^{\{o,d,c\}}$, and from VP_n with $p_n^{\{o,d,c\}}$.

$$Q^{'} \approx spo_n^{\{o,d,c\}} := TS$$

$$TS \begin{cases} sim(s_n^{\{o,d,c\}}, NP_n) = S_n^{sn} \\ sim(o_n^{\{o,d,c\}}, NP_n) = S_n^{on} \\ sim(p_n^{\{o,d,c\}}, VP_n) = S_n^{pv} \\ TS = mean(S_n^{sn}, S_n^{on}, S_n^{pv}) \end{cases}$$

3 Results

We utilized the Trindi Tick list to ascertain the comprehensive functionality of the dialogue engine's ability (Bickmore et al., 2018; Bos et al., 1999). With the exception for the classic Trindi Tick list, according to Bickmore et al. (2018), there is limited formalized standard evaluations for dialogue systems. The Trindi Tick list covers 12 criteria pertaining to handling quality and quantity of user responses, sub-dialogue flows, intelligible responses, etc.

3.1 Is utterance interpretation sensitive to context

No, specifically pertaining to personal contextual information of the user. For example, personal contextual information that is tailored to specific demographic based on behavior or demographic. Another case is a repeat of information that has already been resolved. For example, a user asked a question about side effects of PrEP and then later on the system communicated side effects of PrEP. Ideally it would preferable to skip over the side effect of PrEP if the user had asked the question already.

3.2 Can the system deal with answers to questions that give more information than was requested

Partially yes. This criterion is based on the coverage of the ontology, as the system will attempt to match every predicate that is associated with entities and relationship from a user question.

3.3 Can the system deal with answers to questions that give different information than was actually requested

and

3.4 Can the system deal with answers to questions that give less information than was requested?

Yes, for example, a user can ask a question at any point which will activate the question-answering subroutine. Generally, the system overall is indiscriminate about the amount of information said by the user and will react.

3.5 Can the system deal with ambiguous designators

Partially yes. Similar to criterion 3.2, this is based on the coverage of the ontology and the similarity measurements.

3.6 Can the system deal with negatively specified information

No, not intentionally. We presume that this criterion refers to non-explicit negative responses that compliment the designators, other than stating no or similar negative utterances. With the former, the negated tokens are ignored when the system analyzes the utterance. With the latter, if expected, the system can deal with those responses (e.g., "nah", "negatory", etc.).

> **System**:*PrEP may cause weight loss, serum creatinine increase, headaches, kidney issues, mild nausea, and bone mineral density issues. Following me so far?*
>
> **Participant**: *nope*
>
> **System**: *Hmmm. maybe you have a question. If you do what is the question?*

3.7 Can the system deal with no answer to a question at all

No. The system will wait for a response by the user.

3.8 Can the system deal with noisy input

Yes, by "noisy" we refer to utterances that might be indistinguishable if mis-transcribed or garbled text input.

> **System:** *Great! moving on to the next thing. PrEP may cause weight loss, serum creatinine increase, headaches, kidney issues, mild nausea, and bone mineral density issues. Following me so far?*
>
> **Participant:** *super calif ragilistic expialido cious*

System: *I am sorry. would you mind repeating what you said?*

Participant: *Yes, I follow.*

System: *Great! moving on to the next thing...*

3.9 Can the system deal with "help" sub-dialogs initiated by the user

Yes. A relevant "help" example is when the system facilitates any repeat of information.

System: *headache, mild nausea, and weight loss are mild reactions. Following me so far?*

Participant: *say that again*

System: *headache, mild nausea, and weight loss are mild reactions. Following me so far?*

3.10 Can the system deal with "non-help" sub-dialogs initiated by the user

Yes, a "non-help" sub-dialogue example in our system is whenever a user has additional question that is related to the topic discussed.

System: *PrEP is an antivirual drug, and it also known as Pre-exposure prophylaxis. Following me so far?*

User: *What are some payment options to cover PrEP?*

System: *That is a great question. I would recommend discussing it with your doctor for confirmation. If you have another question, I can help you further. Ask away.* (System did not have an answer and provides one of the default responses.)

User: *no further questions*

System: *Great! moving on to the next thing. PrEP may cause weight loss, serum creatinine increase, headaches, kidney issues, mild nausea, and bone mineral density issues. Following me?*

3.11 Does the system only ask appropriate follow-up questions

Yes. After each segment, the system inquires to the user if they have any follow-up responses to what they heard about PrEP. Example in 3.10 demonstrates this.

3.12 Can the system deal with inconsistent information

Partially yes. The system relies on example of expected utterances to identify the type of utterance using string metric similarity. This may result in misidentifying the utterance and directing the dialogue flow in unintended direction.

4 Discussion

The apparent limitations of the system is highlighted by criterion 3.1, 3.6, and 3.7. The limitation with respect to context is primarily due to lack of a mechanism to handle personalized information. One of the benefits of using ontologies demonstrated by health researchers was the potential to tailor information if we were to capture user information (Bickmore et al., 2011). Previous studies have demonstrated the use of user context ontologies to reason with user data. We assume that this component could be integrated to support personalized contextual information based on group identification or past previous behavior of the user.

Another limitation is the negatively specified information, where if a user were to ask "What if I do not have insurance to pay for PrEP?" Technically the system would not analyze the negative token "not" and focus on the more salient entities of the response. However, a response can be generated by the system, but whether it would accurately respond to the question is unknown, and is determined by the scope of the ontology.

In regards to dealing with no answer, the system awaits for the response of the participant. The reasonable solution is to implement a software code subroutine either on the dialogue system level or on the interface level that times out whenever the user does not provide a timely response. Nonetheless, exploring how this can be done on the ontology-level would need to be investigated and engineered into the ontology.

Other aspects highlighted by our preliminary Trindi Tick assessment underline adherence to criteria regarding handling indistinguishable responses, sub-dialogue branches, and the quality and quantity of information. Also, the quality of the system responsiveness, we theorize, would be dependent on the scope of the knowledge encoded in the ontology.

5 Conclusion

In this paper we present our ontology-based system for handling dialogue for PrEP and PEP counseling. This system also handles questions that are queried from a knowledge base, called the Ontology of PrEP and PEP (OPP). Overall the objective of this work is to demonstrate the feasibility of using an ontology-driven approach to manage automated counseling for PrEP and PEP through a computer-based agent.

Figure 6 shows overall deployment on how the engine will interface with external natural language clients whether they are mobile or terminals (desktops or kiosks). Our eventual goal is to develop a deterministic, planned-based approach within the domain of PrEP and PEP medication adherence (closed domain) and test our approach with live participants.

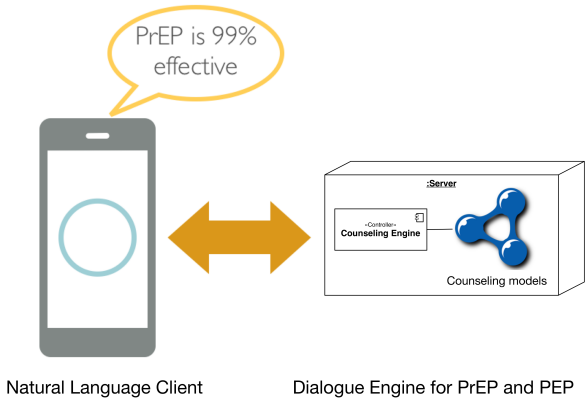

Figure 6: Deployment of a conversational agent client as a mobile application.

Previous research (Amith et al., 2019b) has found limited use of ontologies for medical-based dialogue agents. Solutions addressing PrEP or PEP adherence that researchers have examined include social networks (Kuhns et al., 2017; Garcia et al., 2016) and telehealth solutions (Klausner and CFAR Development Core, 2018; Stekler et al., 2018; Youth Tech Health, 2018). With the former, there has not been any evidence that shows that social networks can address adherence or awareness (Ezennia et al., 2019), and with the latter, telehealth solutions are limited to the availability of a professional and may not be cost effective (Touger and Wood, 2019). Having an automated agent that can provide real-time and high availability to counsel and inform patients may offer an alternative, but further research is needed to foresee this possibility.

Limitations and Future Direction The ontologies that drive the system are currently in draft format and additional work is needed to expand them to include more personalized content, such as where PrEP and PEP can be obtained and information for nonprofit organizations that can provide support, etc. Researchers have conducted simulations to fine tune a formal plan to counsel individuals on the HPV vaccine. Our future work would need to model standard practices for medication counseling adherence that typically happen between patients and providers. This would include conducting simulation studies and working with providers to develop, and then model the counseling flow using the PHIDO framework. Also, from the sample dialogue of the simulation, we can parse out potential questions that can be used to test the question answering component. Lastly, the demonstration of our work is based on text-based modality, and we are working towards interfacing the system to a voice interface to capture the user utterances and evoke the utterance of the machine.

Acknowledgments

Research was supported by the National Library of Medicine of the National Institutes of Health under Award Numbers R01LM011829 and R00LM012104, and the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under Award Number R01AI130460.

References

James Allen. 1995. *Natural language understanding*. Pearson.

American Medical Association. 2016. 8 reasons patients don't take their medications.

Muhammad Amith, Rebecca Lin, Licong Cui, Dennis Wang, Anna Zhu, Grace Xiong, Hua Xu, Kirk Roberts, and Cui Tao. 2019a. An ontology-powered dialogue engine for patient communication of vaccines. In *4th International Workshop on Semantics-Powered Data Mining and Analytics (SEPDA 2019) in Conjunction with the 19th International Semantic Web Conference (ISWC 2019)*.

Muhammad Amith, Rebecca Lin, Rachel Cunningham, Qiwei Luna Wu, Lara S Savas, Yang Gong, Julie A Boom, Lu Tang, and Cui Tao. 2020. Examining potential usability and health beliefs among young adults using a conversational agent for hpv vaccine counseling. In *2020 American Medical Informatics Association Informatics Summit*.

Muhammad Amith, Kirk Roberts, and Cui Tao. 2019b. Conceiving an application ontology to model patient human papillomavirus vaccine counseling for dialogue management. *BMC bioinformatics*, 20(21):1–16.

Aliaksandr Autayeu. 2016. extJWNL.

Timothy Bickmore, Ha Trinh, Reza Asadi, and Stefan Olafsson. 2018. Safety first: Conversational agents for health care. In *Studies in Conversational UX Design*, pages 33–57. Springer.

Timothy W Bickmore, Daniel Schulman, and Candace L Sidner. 2011. A reusable framework for health counseling dialogue systems based on a behavioral medicine ontology. *Journal of biomedical informatics*, 44(2):183–197.

Oni J Blackstock, Brent A Moore, Gail V Berkenblit, Sarah K Calabrese, Chinazo O Cunningham, David A Fiellin, Viraj V Patel, Karran A Phillips, Jeanette M Tetrault, and Minesh Shah. 2017. A cross-sectional online survey of hiv preexposure prophylaxis adoption among primary care physicians. *Journal of general internal medicine*, 32(1):62–70.

Johan Bos, Staffan Larsson, I Lewin, C Matheson, and D Milward. 1999. Survey of existing interactive systems. *Trindi (Task Oriented Instructional Dialogue) report*, (D1):3.

Centers for Disease Control and Prevention. 2014. Preexposure prophylaxis for the prevention of HIV infection in the United States–2014: A clinical practice guideline.

Centers for Disease Control and Prevention. 2016. Monitoring selected national hiv prevention and care objectives by using hiv surveillance data—united states and 6 dependent areas, 2014. Report, Centers for Disease Control and Prevention.

Meredith E Clement, Jessica Seidelman, Jiewei Wu, Kareem Alexis, Kara McGee, N Lance Okeke, Gregory Samsa, and Mehri McKellar. 2018. An educational initiative in response to identified prep prescribing needs among pcps in the southern us. *AIDS care*, 30(5):650–655.

Ogochukwu Ezennia, Angelica Geter, and Dawn K Smith. 2019. The prep care continuum and black men who have sex with men: A scoping review of published data on awareness, uptake, adherence, and retention in prep care. *AIDS and Behavior*, 23(10):2654–2673.

Jonathan Garcia, Caroline Parker, Richard G Parker, Patrick A Wilson, Morgan Philbin, and Jennifer S Hirsch. 2016. Psychosocial implications of homophobia and hiv stigma in social support networks: insights for high-impact hiv prevention among black men who have sex with men. *Health Education and Behavior*, 43(2):217–225.

Birte Glimm, Ian Horrocks, Boris Motik, Giorgos Stoilos, and Zhe Wang. 2014. Hermit: an owl 2 reasoner. *Journal of Automated Reasoning*, 53(3):245–269.

Maja Hadzic, Pornpit Wongthongtham, Tharam Dillon, and Elizabeth Chang. 2009. *Ontology-based multi-agent systems*. Springer.

Kristen L Hess, Shacara D Johnson, Xiaohong Hu, Jianmin Li, Baohua Wu, Chenchen Yu, Hong Zhu, Chang Jin, Mi Chen, and John Gerstle. 2018. Diagnoses of hiv infection in the united states and dependent areas, 2017. *HIV Surveillance Report*.

Daniel Jurafsky and James Martin. 2000. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*, second edition edition. Pearson Prentice Hall, Upper Saddle River, NJ.

Catriona M Kennedy, John Powell, Thomas H Payne, John Ainsworth, Alan Boyd, and Iain Buchan. 2012. Active assistance technology for health-related behavior change: an interdisciplinary review. *Journal of medical Internet research*, 14(3):e80.

J Klausner and CFAR Development Core. 2018. [link].

M.P. Korstanje. 2019. SimMetrics.

Douglas Krakower, Norma Ware, Jennifer A Mitty, Kevin Maloney, and Kenneth H Mayer. 2014. Hiv providers' perceived barriers and facilitators to implementing pre-exposure prophylaxis in care settings: a qualitative study. *AIDS and Behavior*, 18(9):1712–1721.

Lisa M Kuhns, Anna L Hotton, John Schneider, Robert Garofalo, and Kayo Fujimoto. 2017. Use of pre-exposure prophylaxis (prep) in young men who have sex with men is associated with race, sexual risk behavior and peer network size. *AIDS and Behavior*, 21(5):1376–1382.

Christopher D Manning, Mihai Surdeanu, John Bauer, Jenny Rose Finkel, Steven Bethard, and David McClosky. 2014. The stanford corenlp natural language processing toolkit. In *Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations*, pages 55–60.

Alvaro Monge and Charles Elkan. 1997. An efficient domain-independent algorithm for detecting approximately duplicate database records. *Proc. of the ACM-SIGMOD Workshop on Research Issues on Knowledge Discovery and Data Mining*.

Alvaro E Monge, Charles Elkan, et al. 1996. The field matching problem: Algorithms and applications. In *Kdd*, volume 2, pages 267–270.

Rob Shearer, Boris Motik, and Ian Horrocks. 2008. Hermit: A highly-efficient owl reasoner. In *Owled*, volume 432, page 91.

Evren Sirin, Bijan Parsia, Bernardo Cuenca Grau, Aditya Kalyanpur, and Yarden Katz. 2007. Pellet: A practical owl-dl reasoner. *Journal of Web Semantics*, 5(2):51–53.

Dawn K Smith, Jeffrey H Herbst, and Charles E Rose. 2015. Estimating hiv protective effects of method adherence with combinations of preexposure prophylaxis and condom use among african american men who have sex with men. *Sexually transmitted diseases*, 42(2):88–92.

Robert Speer and Joanna Lowry-Duda. 2017. Conceptnet at semeval-2017 task 2: Extending word embeddings with multilingual relational knowledge. *arXiv preprint arXiv:1704.03560*.

Joanne D Stekler, Vanessa McMahan, Lark Ballinger, Luis Viquez, Fred Swanson, Jon Stockton, Beth Crutsinger-Perry, David Kern, and John D Scott. 2018. Hiv pre-exposure prophylaxis prescribing through telehealth. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 77(5):e40–e42.

Rebecca Touger and Brian R Wood. 2019. A review of telehealth innovations for hiv pre-exposure prophylaxis (prep). *Current HIV/AIDS Reports*, 16(1):113–119.

Dominic Widdows and Trevor Cohen. 2010. The semantic vectors package: New algorithms and public tools for distributional semantics. In *2010 IEEE Fourth International Conference on Semantic Computing*, pages 9–15. IEEE.

Brian R Wood, Vanessa M McMahan, Kelly Naismith, Jonathan B Stockton, Lori A Delaney, and Joanne D Stekler. 2018. Knowledge, practices, and barriers to hiv preexposure prophylaxis prescribing among washington state medical providers. *Sexually transmitted diseases*, 45(7):452–458.

Michael Wooldridge. 2009. *An introduction to multiagent systems*. John Wiley and Sons.

Michael Wooldridge and NR Jennings. 1995. Intelligent agents: Theory and practice the knowledge engineering review. *The knowledge engineering review*, 10:115–152.

World Health Organization. 2017. World health statistics 2017: Monitoring health for the sdgs.

Youth Tech Health. 2018. [link].