

Summer 8-14-2019

## Design and Architecture of an Ontology-driven Dialogue System for HPV Vaccine Counseling

Amith Faheem Muhammad

Follow this and additional works at: [https://digitalcommons.library.tmc.edu/uthshis\\_dissertations](https://digitalcommons.library.tmc.edu/uthshis_dissertations)



Part of the [Bioinformatics Commons](#), and the [Medicine and Health Sciences Commons](#)

---

# Design and Architecture of an Ontology-driven Dialogue System for HPV Vaccine Counseling

A  
DISSERTATION

PRESENTED TO THE FACULTY OF  
UNIVERSITY OF TEXAS HEALTH SCIENCE CENTER AT HOUSTON  
SCHOOL OF BIOMEDICAL INFORMATICS  
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

By

MUHAMMAD “TUAN” AMITH, MS

UNIVERSITY OF TEXAS HEALTH SCIENCE CENTER AT HOUSTON

2019

*Dissertation Committee:*

Cui Tao, PhD<sup>1</sup>, Advisor

Yang Gong, MD, PhD<sup>1</sup>

Hua Xu, PhD<sup>1</sup>

Kirk Roberts, PhD<sup>1</sup>

Julie A. Boom, MD<sup>2,3</sup>

<sup>1</sup>School of Biomedical Informatics,  
University of Texas Health Science Center at Houston

<sup>2</sup>Texas Children’s Hospital

<sup>3</sup>Baylor College of Medicine

created in  
X<sub>Y</sub>LaTeX



©2019 – MUHAMMAD “TUAN” AMITH, MS  
ALL RIGHTS RESERVED.

---

DEDICATED TO MY LATE FATHER (“BAPEE”, I.E. BAPAK), WHO TAUGHT ME EVERYTHING HE KNEW ABOUT ENGINEERING, COMPUTERS, SCI-FI, BASKETBALL, AND THE UNVARNISHED TRUTHS OF HUMAN NATURE AND ON HOW TO SURVIVE IN THIS WILD WORLD.

MOST OF ALL, I APPRECIATE HIM SHOWING ME THAT CHARACTER, INTEGRITY, HONESTY, CHARITY, AND HEART COUNTS FOR SOMETHING. I MISS YOU ALWAYS, FAA.

# Acknowledgments

AFTER MANY YEARS, I learned the hard way that no man is an island, and this dissertation would not have been possible without the kind support of everyone that I know — be it a piece of advice, a critique, a helping hand, etc. A little can go a long way. For that, I am grateful for all who helped on this journey.

I FEEL PRIVILEGED to work under the tutelage of the best faculty members that the School of Biomedical Informatics (SBMI) has to offer to serve on my dissertation committee. Most of all, I thank Dr. Cui Tao for being my supervisor, and her support through the ups and downs of my dissertation. As one of the best biomedical informatics researchers in the field, I am honored to be one of her students. I thank Dr. Yang Gong, Dr. Hua Xu, and Dr. Kirk Roberts for their renowned informatics expertise, input, and support of my research. Also, I want to thank Dr. Julie A. Boom, MD, for her pediatric and public health expertise that was greatly needed to shape this project. Last but not least, I thank Dr. Craig Johnson who helped guide me with the early statistical knowledge and various career and dissertation advice.

IN HINDSIGHT, when I compare what we have done now to what we started, the current evolution of the project would not have been possible without the above-mentioned individuals.

ALL IN ALL, my successes are due to them, and my failures are my own.

SPECIAL THANK YOU to my friends and colleagues at SBMI for their emotional support and sharing their time and research knowledge – Chen Liang, Judy Song & Patrick Sun, Jingcheng Du, Frank Manion, Hong Kang, Xinyuan Zhang, and all the bright minds that are part of Dr. Tao's SBMI Ontology Research Group (SORG).

IT HAS BEEN A PLEASURE collaborating with smart and talented summer undergraduates – Rebecca Lin, Dennis Wang, and Anna Zhu. Thank you for bringing your “A” game to the research project, and I wish you all the best in your future careers.

EVERY JOURNEY has a cast of supporting characters....

I WAS VERY FORTUNATE to have the Dean of SBMI, Jaijie Zhang, PhD, as a professor for an early course. Dr. Zhang's help was instrumental in encouraging me to succeed in the beginning, and he has been a valued supporter ever since. Earlier, Dr. Zhang gave one important advice for a successful

dissertation that helped in the development of my dissertation. Also, I thank Dr. Elmer Bernstam who gave the best advice in order to succeed at a PhD career, which later catapulted me to where I am now. Also, thank you to Rachel Cunningham, MPH, for providing the additional guidance on the public health contribution of this work (and also for her patience with me).

THANK YOU to the UTHealth Innovation for Cancer Prevention Research Training Program faculty — Patricia A. Mullen, Roberta Ness, David Loose, and Lara S. Savas — for providing support and their own insight into my research.

ALSO THANK YOU to the following for their assistance and intellectual advice throughout the years: Sivaram Arabandi, Abraham Bernstein, Yong Chen, Trevor Cohen, Licong Cui, Kayo Fujimoto, Yolanda Gil, Craig Harrington, Marcelline Harris, Yongqun He, Zhe He, Tiffany Champagne-Langabeer, James Lewis (IBM), Paul G. Loubser, Tang Lu, Frank Manion, Natasha Noy, Lea Sacca, Roger Schavendeldt, Laura A. Shay, Dean Sittig, Lena Wu, Yaoyun Zhang, and Min Zhu.

RESEARCH FUNDING for this project over the last four years was provided through the UTHealth Innovation for Cancer Prevention Research Training Program (Cancer Prevention and Research Institute of Texas grant # RPI60015), and through various NIH grants documented in published works.

## Design and Architecture of an Ontology-driven Dialogue System for HPV Vaccine Counseling

### ABSTRACT

Speech and conversational technologies are increasingly being used by consumers, with the inevitability that one day they will be integrated in health care. Where this technology could be of service is in patient-provider communication, specifically for communicating the risks and benefits of vaccines. Human papillomavirus (HPV) vaccine, in particular, is a vaccine that inoculates individuals from certain HPV viruses responsible for adulthood cancers - cervical, head and neck cancers, etc. My research focuses on the architecture and development of speech-enabled conversational agent that relies on series of consumer-centric health ontologies and the technology that utilizes these ontologies. Ontologies are computable artifacts that encode and structure domain knowledge that can be utilized by machines to provide high level capabilities, such as reasoning and sharing information. I will focus the agent’s impact on the HPV vaccine domain to observe if users would respond favorably towards conversational agents and the possible impact of the agent on their beliefs of the HPV vaccine. The approach of this study involves a multi-tier structure. The first tier is the domain knowledge base, the second is the application interaction design tier, and the third is the feasibility assessment of the participants. The research in this study proposes the following questions:

1. *Can ontologies support the system architecture for a spoken conversational agent for HPV vaccine counseling?*
2. *How would prospective users’ perception towards an agent and towards the HPV vaccine be impacted after using conversational agent for HPV vaccine education?*

Dissertation advisor: Cui Tao, PhD

Muhammad “Tuan” Amith, MS

The outcome of this study is a comprehensive assessment of a system architecture of a conversational agent for patient-centric HPV vaccine counseling. Each layer of the agent architecture is regulated through domain and application ontologies, and supported by the various ontology-driven software components that I developed to compose the agent architecture. Also discussed in this work, I present preliminary evidence of high usability of the agent and improvement of the users’ health beliefs toward the HPV vaccine. All in all, I introduce a comprehensive and feasible model for the design and development of an open-sourced, ontology-driven conversational agent for any health consumer domain, and corroborate the viability of a conversational agent as a health intervention tool.

# Vita

2003 ..... B.S., Computer Science (cum laude), University of Houston-Victoria  
2010 ..... M.S., Software Engineering, University of Houston-Clear Lake  
2014–2015 ..... Graduate Research Assistant, University of Texas Health Science Center at Houston  
2015–2019 ..... Predoctoral Fellow, University of Texas Health Science Center at Houston

## PUBLICATIONS

\* co-first authors

Amith, M., Lin, R., Cui, L., Wang, D., Zhu, A., Xiong, G., Xu, H., Roberts, K., & Tao, C. (2019).

*An Ontology-Powered Dialogue Engine For Patient Communication of Vaccines*. 19th International Semantic Web Conference (ISWC 2019). (Accepted)

Amith, M., Manion, F., Liang, C., Harris, M., Wang, D., He, Y., & Tao, C. (2019). *Architecture*

*and Usability of OntoKeeper, an Ontology Evaluation Tool*. BMC Medical Informatics and Decision Making. (In press)

Amith, M., Fujimoto, K., & Tao, C. (2019). *NET-EXPO: A Gephi Plugin Towards Social Network*

*Analysis of Network Exposure for Unipartite and Bipartite Graphs*. In International Conference on Human-Computer Interaction (pp. 3-12). Springer, Cham

Imahashi, M., Fujimoto, K., Kuhns, L. M., Amith, M., & Schneider, J. A. (2019). *Network overlap*

*and knowledge of a partner's HIV status among young men who have sex with men*. AIDS care, 1-7.

Amith, M., Zhu, A., Cunningham, R., Lin, R., Savas, L., Shay, L., Boom, J., Chen, Y., Gong, Y.,

- Roberts, K., & Tao, C. (2019). *Early Usability Assessment of a Conversational Agent for HPV Vaccination*. Studies in health technology and informatics, 257, 17-23. PMID:30741166. doi.org/10.3233/978-1-61499-951-5-17.
- Fujimoto, K., Wang, P., Flash, C. A., Kuhns, L. M., Zhao, Y., Amith, M., & Schneider, J. A. (2019). *Network modeling of PrEP uptake on referral networks and health venue utilization among young men who have sex with men*. AIDS and Behavior, 23(7), 1698-1707.
- Amith, M., Manion, F., Liang, C., Harris, M., Wang, D., He, Y., & Tao, C. (2018). *OntoKeeper: Semiotic-driven Ontology Evaluation Tool For Biomedical Ontologists*. In 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (pp. 1614-1617). IEEE.
- Amith, M., & Tao, C. (2018). *Representing vaccine misinformation using ontologies*. Journal of Biomedical Semantics, 9(1), 22. PMID:30170633 doi.org/10.1186/s13326-018-0190-0
- Lin, R., Amith, M., Liang, C., Duan, R., Chen, Y., & Tao, C. (2018). *Visualized Emotion Ontology: a model for representing visual cues of emotions*. BMC medical informatics and decision making, 18(2), 64.
- Amith, M., Lin, R., Liang, C., Gong, Y., & Tao, C. (2018). *VEO-Engine: Interfacing and Reasoning with an Emotion Ontology for Device Visual Expression*. In International Conference on Human-Computer Interaction (pp. 349-355). Springer, Cham.
- Amith, M., He, Z., Bian, J., Lossio-Ventura, J. A., & Tao, C. (2018). *Assessing the practice of biomedical ontology evaluation: Gaps and opportunities*. Journal of Biomedical Informatics, 80, 1-13.
- Lin, R., Amith, M., Liang, C., & Tao, C. (2017). *Designing an ontology for emotion-driven visual representations*. In 2017 IEEE International Conference on Bioinformatics and Biomedicine



(BIBM) (pp. 1280-1283). IEEE.

Amith, M., Cunningham, R., Savas, L. S., Boom, J., Schvaneveldt, R., Tao, C., & Cohen, T. (2017).

*Using Pathfinder networks to discover alignment between expert and consumer conceptual knowledge from online vaccine content.* Journal of Biomedical Informatics, 74, 33-45.

Amith, M.\*, Manion, F. J.\*, Harris, M. R.\*, Zhang, Y., Xu, H., & Tao, C. (2017). *Expressing biomedical ontologies in natural language for expert evaluation.* Studies in health technology and informatics, 245, 838.

Amith, M.\*, Song, H. Y.\*, Zhang, Y., Xu, H., & Tao, C. (2017). *Lightweight predicate extraction for patient-level cancer information and ontology development.* BMC medical informatics and decision making, 17(2), 73.

Amith, M., Zhang, Y., Xu, H., & Tao, C. (2017). *Knowledge-Based Approach for Named Entity Recognition in Biomedical Literature: A Use Case in Biomedical Software Identification.* In International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems (pp. 386-395). Springer, Cham.

Amith, M., & Tao, C. (2017). *Modulated evaluation metrics for drug-based ontologies.* Journal of biomedical semantics, 8(1), 17.

Amith, M. (2016). *Ontology-Based Dialogue Systems for Improved Patient HPV Vaccine Knowledge and Perception.* In Doctoral Consortium at International Semantic Web Conference, in: Proceedings of the Doctoral Consortium at the 15th International Semantic Web Conference (ISWC 2016) (pp. 9-16).

Wang, D., Cunningham, R. M., Boom, J., Amith, M., & Tao, C. (2016). *Towards a HPV Vaccine*

*Knowledgebase For Patient Education Content*. Studies in Health Technology and Informatics, 225 (pp. 432–436).

Amith, M., & Tao, C. (2015). *A web application towards semiotic-based evaluation of biomedical ontologies*. In International Workshop on Biomedical Data Mining, Modeling, and Semantic Integration: A Promising Approach to Solving Unmet Medical Needs, BDM2I 2015-co-located with the 14th International Semantic Web Conference, ISWC 2015. CEUR-WS.

Myneni, S., Amith, M., Geng, Y., & Tao, C. (2015). *Towards an ontology-driven framework to enable development of personalized mHealth solutions for Cancer survivors' engagement in healthy living*. Studies in health technology and informatics, 216, 113. *MedInfo Best Paper Finalist*.

Amith, M., Gong, Y., Cunningham, R., Boom, J., & Tao, C. (2015). *Developing VISO: vaccine information statement ontology for patient education*. Journal of biomedical semantics, 6(1), 23.

Amith, M., & Tao, C. (2014). *Toward the Design of a Vaccine Information Ontology for Patient Education*. International Workshop on Vaccine and Drug Ontology Studies (VDOS), International Conference on Biomedical Ontology (ICBO) 2014.

Amith M., & Tao C. (2014) *A Prototype Mobile Virtual Assistant for Semantic-Based Vaccine Information Retrieval*. In: Zheng X., Zeng D., Chen H., Zhang Y., Xing C., Neill D.B. (eds) Smart Health. ICSH 2014. Lecture Notes in Computer Science, vol 8549. Springer, Cham

Amith, M., Loubser, P. G., Chapman, J., Zoker, K. C., & Ferreira, F. E. R. (2012). *Optimization of an EHR mobile application using the UFuRT conceptual framework*. In AMIA Annual Symposium Proceedings (Vol. 2012, p. 209). American Medical Informatics Association.

White, S. & Amith, M. (2012). *Software Design: A 2011 Primer*. ACET Journal of Computer Education and Research 7.

*In Review and Revision*

Amith, M., Cohen, T., Cunningham, R., Smith, N., Savas, L., Cuccaro, P., Boom, J., Schvaneveldt, R., & Tao, C. *Mining HPV Vaccine Knowledge Structures of Young Adults from Reddit Using Distributional Semantics and Pathfinder Networks*. Cancer Control Journal.

Amith, M., Roberts, K., & Tao, C. *Conceiving an application ontology to model patient human papillomavirus vaccine counseling for dialogue management*. BMC Bioinformatics.

Amith, M., Fujimoto, K., Mauldin, R., & Tao, C. *Friend of a Friend with Benefits Ontology (FOAF+): Extending a social network ontology for public health*. Journal of Biomedical Semantics.

Wei, Q., Zhang, Y., Amith, M., Lin, R., Lapeyrolerie, J., Tao, C., & Xu, H. *Recognizing Software Names in Biomedical Literature Using Machine Learning*. Health Informatics Journal.

*In Preparation*

Mauldin, R., Amith, M., Kuhns, L., Schneider, J., & Fujimoto, K. *Racial/ethnic disparity in PrEP care continuum: PrEP referral networks involving health venues and younger MSM*.

ABSTRACTS

Amith, M., Fujimoto, K., & Tao, C. (2019). *FOAF+: Friend of a Friend with Benefits Ontology*. XXXIX Sunbelt Social Networks Conference of the International Network for Social Network Analysis (INSNA).

- Bhadriraju, M., & Amith, M. (2019). *Social Media as a Medium for Targeted HIV/AIDS Intervention in Youth MSM Populations*. Information Technology and Communications in Health (ITCH 2019).
- Zhu, A., Amith, M., Lu, M., & Tao, C. (2019). *Vaccination Vacation: A Story-Based Game to Improve HPV Vaccine Knowledge and Awareness*. Information Technology and Communications in Health (ITCH 2019).
- Manion, F. J., Harris, M. R., Amith, M., & Tao, C. (2018) *Development and Validation of an Ontological Representation of the US Common Rule*. American Medical Informatics Association 2018 Annual Symposium.
- Amith, M., Cuccaro, P., & Savas, L. (2017). *Preliminary Textual Big Data Analysis of HPV-related Discourse on Reddit*. 2017 CPRIT Conference, Austin, TX, November, 2017.
- Lin, R., Amith, M., & Tao, C. (2017). *A Model for Emotion-based Visualizations for Conversational Agents in HPV Vaccine Counseling*. 2017 CPRIT Conference, Austin, TX, November, 2017.
- Manion, F., Harris, M., Song, H., He, Y., Amith, M., Xu, H., & Tao, C. (2017). *From Scoping Review to Metadata: An Evidence-Based Approach to Identifying Informed Consent Metadata for Biorepositories*. American Medical Informatics Association 2017 Annual Symposium, Washington, DC, November, 2017.
- Xu, H., Grethe, J.S., [et al, including Amith, M.]. (2016). *Development of DataMed, a Data Discovery Index Prototype by bioCADDIE: Laying the Ground- work for Biomedical Data Discovery*. American Medical Informatics Association 2016 Annual Symposium, Chicago, IL, November, 2017

- Grethe, J.S., Ozyurt, B., [et al, including Amith, M.]. 2016. *A Scalable Dataset Indexing Infrastructure for the bioCADDIE Data Discovery System*. American Medical Informatics Association 2016 Annual Symposium, Chicago, IL, November 2017.
- Wang, D., Amith, M., & Tao, C. (2015). *HPV VISO: Encoding HPV Vaccine Education Content For Patient Knowledge-based Systems*. 2015 CPRIT Conference, Austin, TX, November 2015.
- Mays, M., Amith, M., Garcia, E., Pacheco, P., & Johnson, C. (2014). *Predictors of Perceived Health Status Among Low Income Community Residents in Puerto Rico*. Seventh Health Disparities Conference, New Orleans, LA, March 2014.
- White, S., & Amith, M. (2010). *Software Design - 2010 Primer*. Proceedings of ACET 46th Annual Conference, League City, TX, October 2010.

#### GRANTS AND AWARDS

- Student Travel Award, Sunbelt Social Networks Conference of the International Network for Social Network Analysis (INSNA), 2019.
- Willerson Endowed Scholarship Award, University of Texas Health Science Center, 2018.
- Dr. Aoki Travel Award, School of Biomedical Informatics, 2018.
- Student Travel Scholarship, National Science Foundation, International Conference Biomedical Ontologies, Corvallis, Oregon, 2018.
- Thomas F. Burks Scholarship, University of Texas Health Science Center, 2018.
- Willerson Endowed Scholarship Award, University of Texas Health Science Center, 2017.

Student & Trainee Award, National Science Foundation, International Conference on Intelligent Biology and Medicine, Houston, TX, 2016.

International Semantic Web Conference Travel Grant, National Science Foundation, International Semantic Web Conference, Kobe, Japan, 2016.

Dr. Aoki Travel Award, School of Biomedical Informatics, 2016. BioCreative/International Conference Biomedical Ontologies Fellowship, National Science Foundation, International Conference Biomedical Ontologies, Corvallis, Oregon, 2016.

Merrill and Joe Hafner Doctoral Award, School of Biomedical Informatics, 2016.

Dr. Aoki Travel Award, School of Biomedical Informatics, 2015.

MedInfo Best Paper Finalist for *Towards an Ontology-driven Framework to Enable Development of Personalized mHealth Solutions for Cancer Survivors' Engagement in Healthy Living*, Medinfo 2015, Sao Paolo, Brazil, 2015.

International Semantic Web Conference Travel Grant, National Science Foundation, International Semantic Web Conference, Bethlehem, PA, 2015.

Best of mHealth for *Optimization of an EHR Mobile Application Using the UfuRT Conceptual Framework*, American Medical Informatics Association, Chicago, IL, 2012.

FIELD OF STUDY  
Biomedical Informatics

# Contents

1	INTRODUCTION	1
1.1	Objective . . . . .	7
1.2	Research Hypothesis and Specific Aims . . . . .	8
1.3	Dissertation Summary . . . . .	9
2	BACKGROUND	12
2.1	Human Papillomavirus Vaccine . . . . .	13
2.1.1	HPV Vaccine & Population Impact . . . . .	15
2.1.2	Causes of Low HPV Vaccination Coverage . . . . .	17
2.1.3	Directions for Increased HPV Vaccination Uptake . . . . .	22
2.1.4	Health Belief Model . . . . .	25
2.2	Ontologies . . . . .	26
2.2.1	Ontology Evaluation . . . . .	31
2.2.2	Ontology-based Question-Answering . . . . .	34
2.3	Natural Language Processing . . . . .	36
2.3.1	Open Information Extraction . . . . .	37
2.3.2	Natural Language Generation . . . . .	37
2.3.3	Dialogue Systems . . . . .	38
2.3.4	Ontology-based Dialogue Systems for Health . . . . .	42
2.4	Agents and Agent Architecture . . . . .	45
2.4.1	Ontologies for Agent Architecture . . . . .	48
2.5	Summary . . . . .	49
3	HEALTH CONSUMER VACCINE ONTOLOGIES	50
3.1	Ontology Evaluation Tools and Methods . . . . .	51
3.1.1	Natural language generation with Hootation . . . . .	52
3.1.2	Benchmarking ontologies with semiotic metric suite . . . . .	55
3.1.3	OntoKeeper . . . . .	56
3.2	Domain Knowledge Base for Vaccines . . . . .	59
3.2.1	Vaccine Information Statement Ontology . . . . .	60
3.2.2	Vaccine Misinformation Ontology . . . . .	65
3.2.3	Anti-vaccination Information concepts . . . . .	68
3.2.4	Controversial Vaccine concepts . . . . .	69
3.2.5	Information Theory concepts . . . . .	70

3.2.6	Cues to motivation and indication concepts . . . . .	75
3.2.7	Vaccine Theories concepts . . . . .	77
3.2.8	VAXMO's Metrics . . . . .	80
3.3	Summary . . . . .	81
4	INTERACTION-DRIVEN APPLICATION ONTOLOGIES . . . . .	82
4.1	Wizard of OZ experiment . . . . .	83
4.1.1	Tools and Methods . . . . .	85
4.2	Ontology-driven dialogue engine . . . . .	89
4.2.1	Patient Health Information Dialogue Ontology . . . . .	91
4.2.2	Conversational Ontology Operator Engine (COO) . . . . .	102
4.3	Ontology-driven visual interface for agents . . . . .	111
4.3.1	Visualized Emotion Ontology . . . . .	113
4.3.2	VEO-Engine . . . . .	119
4.4	Summary . . . . .	122
5	ARCHITECTURAL ASSESSMENT AND ANALYSIS OF COMPONENTS . . . . .	123
5.1	Evaluation of the domain ontology . . . . .	125
5.1.1	Vaccine Misinformation Ontology quality assessment . . . . .	126
5.1.2	VISO-HPV quality assessment . . . . .	128
5.1.3	Visualized Emotion Ontology (VEO) quality assessment . . . . .	129
5.1.4	Patient Health Information Dialogue Ontology (PHIDO) assessment . . . . .	132
5.2	Application of ontologies for dialogue system components . . . . .	134
5.2.1	PHIDO for managing the dialogue engine (COO) . . . . .	135
5.2.2	Question answering (FOQUS) for VISO-HPV . . . . .	144
5.2.3	VEO-Engine . . . . .	147
5.3	Examining usability and impact on the users' health beliefs . . . . .	149
5.3.1	Early Usability Assessment of Conversational Agent . . . . .	149
5.3.2	Extended Usability Assessment of the Conversational Agent . . . . .	152
5.4	Summary . . . . .	158
6	CONCLUSION . . . . .	160
6.1	Discussion . . . . .	161
6.1.1	User Experience . . . . .	162
6.1.2	Interface layer . . . . .	166
6.1.3	Application layer . . . . .	168
6.1.4	Domain layer . . . . .	171
6.2	Contribution . . . . .	174
6.3	Future Direction and Limitations . . . . .	177



APPENDIX A	SAMPLE ONTOLOGIES FROM NATIONAL CENTER FOR BIOMEDICAL ONTOLOGY BIOPORTAL	181
APPENDIX B	VACCINE INFORMATION STATEMENT ONTOLOGY	183
APPENDIX C	DEFINITION OF EMOTIONS FROM THE VISUALIZED EMOTION ONTOLOGY	187
APPENDIX D	FRANKENSTEIN ONTOLOGY QUESTION-ANSWERING FOR USER-CENTRIC SYSTEM	190
APPENDIX E	USABILITY VARIABLES	192
REFERENCES		196

# Listing of Tables

Table 2.1	Papers surveyed for ontology evaluation software tools. . . . .	32
Table 3.1	Sample output showing axiom data from ICO. . . . .	54
Table 3.2	Evaluators results from examining natural language output. . . . .	54
Table 3.3	NCBO Sample Aggregate Scores . . . . .	56
Table 4.1	System Utterance classes . . . . .	93
Table 4.2	Participant Utterance classes . . . . .	94
Table 4.3	Outline list of the various Speech Tasks and their function/purpose. . . . .	95
Table 4.4	Visualization motifs for emotions. . . . .	117
Table 5.1	Comparison of quality scoring derived from semiotic metric suite for VAXMO and the NCBO BioPortal sample. . . . .	126
Table 5.2	Comparison of quality scoring derived from semiotic metric suite for VISO-HPV and the NCBO BioPortal sample. . . . .	128
Table 5.3	Quality scores comparing the VEO with cognitive ontologies. . . . .	130
Table 5.4	Survey results of visualization . . . . .	131
Table 5.5	Comparison of quality scoring derived from semiotic metric suite for PHIDO and the NCBO BioPortal sample. . . . .	133
Table 5.6	Agreement ratings for the question answering component. . . . .	146
Table 5.7	Accuracy of the question answering component. . . . .	146
Table 5.8	Comparing CHAIS constructs from the intervention and published studies. .	158
Table A.1	Popular ontologies from NCBO on September 2015. . . . .	182
Table B.1	Vaccine Information Statement Ontology Class Table . . . . .	184
Table B.2	Vaccine Information Statement Ontology's object properties . . . . .	185
Table C.1	Definition of positive emotions. . . . .	188
Table C.2	Definition of negative emotions . . . . .	189
Table E.1	Participant demographic variables . . . . .	193
Table E.2	System Usability Survey variables . . . . .	193
Table E.3	Speech User Interface Service Quality survey . . . . .	194
Table E.4	CHIAS health belief model (HBM) . . . . .	195

# Listing of Figures

Figure 1.1	Linked open data cloud . . . . .	4
Figure 1.2	DIKW (Data, Information, Knowledge, Wisdom) pyramid . . . . .	6
Figure 1.3	Application of knowledge pyramid for agents. . . . .	6
Figure 1.4	Development for the architecture of HPV vaccine conversational agent . . . .	8
Figure 2.1	Adapted figure from Janz & Becker (1984) for HPV vaccine intervention. . .	26
Figure 2.2	Example of ontology abstraction. . . . .	28
Figure 2.3	Simple implementation of dialogue management using finite state automaton. .	40
Figure 2.4	Basic architecture of a dialogue system . . . . .	42
Figure 3.1	Login screen for OntoKeeper. . . . .	57
Figure 3.2	One of four panels (Semantic score) in OntoKeeper . . . . .	58
Figure 3.3	Panel for the aggregated score . . . . .	58
Figure 3.4	Process of creating models from each VIS document to create a singular common knowledge structure model. . . . .	61
Figure 3.5	Dosage Pattern from VISO. . . . .	62
Figure 3.6	Reaction Pattern from VISO. . . . .	63
Figure 3.7	Target Pattern from VISO. . . . .	63
Figure 3.8	Screenshot VISO-HPV in Protégé. . . . .	65
Figure 3.9	Class-level structure of the Vaccine Misinformation Ontology (VAXMO) . .	66
Figure 3.10	Class structure of VAXMO showing the Misinformation concept and it associated classes, like Anti-vaccination Information. . . . .	67
Figure 3.11	Class structure of VAXMO showing the Subject concept and its associated classes, like Contraversial_Vaccine . . . . .	70
Figure 3.12	Class structure of VAXMO showing concepts related to Information Theory process, which were derived from the original Misinformation Ontology. . . . .	71
Figure 3.13	Class structure of VAXMO showing Communication Channel, a concept from the original Misinformation Ontology. . . . .	74
Figure 3.14	Class structure of VAXMO showing the Motivation concept and also the Indication class. . . . .	76
Figure 3.15	Class structure of VAXMO showing the False Asserted Nanopublication concept that is a subclass of Publication. . . . .	79
Figure 3.16	Details of the of the False Asserted Nanopublication class. . . . .	80
Figure 4.1	Wizard of OZ experimental protocol. . . . .	84

Figure 4.2	Simulation software for the study's Wizard of OZ experiment. . . . .	85
Figure 4.3	UML component diagram of Wizard of OZ software . . . . .	87
Figure 4.4	General framework for discussing vaccine. . . . .	88
Figure 4.5	The Salutation class from PHIDO. . . . .	96
Figure 4.6	The Valediction class from PHIDO. . . . .	96
Figure 4.7	The Initiate Discussion class from PHIDO. . . . .	97
Figure 4.8	This Transition to Topic class from PHIDO. . . . .	97
Figure 4.9	The Interview Participant class from PHIDO. . . . .	98
Figure 4.10	The Discuss Health Topic class from PHIDO. . . . .	98
Figure 4.11	The Question Answering Task class from PHIDO. . . . .	99
Figure 4.12	UML class diagram of the relationship between the Communication Goal class with Speech Task class and the Discussion class. . . . .	101
Figure 4.13	Screenshot of Protege with PHIDO. . . . .	101
Figure 4.14	A UML sequence diagram describing the transition mechanism for PHIDO's dialogue flow . . . . .	103
Figure 4.15	Extraction process from ontology for FOQUS. . . . .	105
Figure 4.16	Question analysis for FOQUS . . . . .	106
Figure 4.17	Similarity scoring between question data and ontology data for FOQUS . . . . .	107
Figure 4.18	Selection of answer triples based on ranks for FOQUS . . . . .	108
Figure 4.19	Examples of virtual agents. . . . .	III
Figure 4.20	Leonardo da Vinci: Diagram of a proposed flying machine (1789) . . . . .	II2
Figure 4.21	Jackson Pollock from MoMA NYC. . . . .	II3
Figure 4.22	Graphical structure of the revised OCC model that includes the emotion surprise from Ekman . . . . .	II5
Figure 4.23	Graphical structure of revised OCC model showing each emotion associated with a composite visualization . . . . .	II6
Figure 4.24	Graphical representation of the Emotion concept from VEO. . . . .	II8
Figure 4.25	Graphical representation of the Visualization concept from VEO. . . . .	II8
Figure 4.26	Mapping of the knowledge hierarchy as the process for machine interpretation of emotions. . . . .	120
Figure 4.27	Protégé screenshot of the relief visualization with a link to an image on the Web and a local image. . . . .	121
Figure 5.1	Modified version diagram of the Discuss Health Task with the Transition Topic utterance. . . . .	136
Figure 5.2	Screenshot of Protégé with PHIDO's instances for testing. . . . .	137
Figure 5.3	Dialogue interaction showing confirming health information . . . . .	137
Figure 5.4	Dialogue interaction showing disconfirming health information . . . . .	138
Figure 5.5	Dialogue interaction showing requesting the repeat of health information . . . . .	139
Figure 5.6	Dialogue interaction facilitating misunderstood utterances from the user . . . . .	139

Figure 5.7	Dialogue interaction showing the transition from health information exchange to question answering mode (simulated) . . . . .	140
Figure 5.8	An alternate dialogue interaction showing the transition from health information exchange to question answering mode (simulated) . . . . .	141
Figure 5.9	Dialogue interaction showing a question answered . . . . .	142
Figure 5.10	Dialogue interaction showing a question with no answer . . . . .	143
Figure 5.11	Dialogue interaction providing the user the option to ask another question . . . . .	144
Figure 5.12	Touchscreen device displaying results of a visualization query for the emotion <i>anger</i> . . . . .	148
Figure 5.13	VEO-Engine performed a reasoning task based on parameters for the emotion of <i>hope</i> . . . . .	148
Figure 5.14	SUS Scores among different vaccination statuses. . . . .	156
Figure 6.1	Layered architecture of the conversational agent for HPV vaccine. . . . .	161
Figure 6.2	Interpreting the user's emotional information. . . . .	171
Figure 6.3	Hardware deployment for a conversational agent . . . . .	175
Figure B.1	TBox level structure of VISO. . . . .	186
Figure D.1	Process diagram of Frankenstein Ontology Question-Answering for User-Centric System . . . . .	191

# Listing of Code

Listing 3.1	Sample syntax coding from Time Event Ontology in Turtle syntax . . . . .	52
Listing 3.2	Turtle syntax of Misinformation and Anti-vaccine Information from VAXMO	68
Listing 3.3	Turtle syntax of Misinformation and Anti-vaccine Information from VAXMO	70
Listing 3.4	Turtle syntax of Participants and Familiarity concepts from VAXMO . . . . .	71
Listing 3.5	Turtle syntax of Communication Channel, Synchronicity, Modality, Distribution Method, and Recordability concepts from VAXMO . . . . .	74
Listing 3.6	Turtle syntax of Motivation and Indication from VAXMO . . . . .	76
Listing 3.7	Sample nanopublication encoding adapted from Groth et al. (2013) . . . . .	78
Listing 4.1	SPARQL query for a visualization from VEO . . . . .	121

*The scariest moment is always just before you start.*

Stephen King

# 1

## Introduction

FROM THE TIME WE ARE BORN, we learn to interact with one another using speech to achieve a variety of human objectives – request for food or milk, attention from another human, or to learn. It is natural for human beings to use speech and dialogue interaction to communicate and exchange information that could improve the quality of our lives. The research contained in this dissertation embarks on an attempt at designing a conversational agent using a series of ontological knowledge bases that provide the intelligence of the agent to interact with human users. The specific use-case is a light counseling session discussing the HPV vaccine with a patient (i.e. health consumer) that should occur between patient and provider. This proof of concept conversational agent could be integrated in a waiting room area of a clinic where they are free to be engaged or waiting for a

clinician.

Several researchers have stated the importance of vaccines to diminish the impact of death-related disease. The HPV vaccine is no different. This particular vaccine protects individuals from some cancer-causing viruses like throat cancer and cervical cancer. The human papillomavirus (HPV) vaccine is noted to be 99% effective against the HPV virus ([Centers for Disease Control and Prevention, 2014a](#)), and yet the United States population uptake for HPV vaccine is at a low 37.1% ([Walker et al., 2017](#)), short of the 80% uptake goal ([US Department of Health and Human Services and Office of Disease Prevention and Health Promotion and others, 2012](#)). Many life threatening cancers, like cervical cancer, is attributed to high risk HPV viruses (Type 16 and 18). One predictive study reported that if the population were to have 70% coverage of HPV vaccine, at least 4 million deaths could be prevented in the next decade ([Goldie et al., 2008](#)). Misinformation relating to the safety and importance of vaccines affects coverage rates of all vaccines, including but not limited to the HPV vaccine ([Myers & Pineda, 2009](#)). Closing this misinformation gap is crucial if we are to guarantee the safety of the human population from viruses and diseases, and from diseases long-thought to be extinguished from re-emerging in the community.

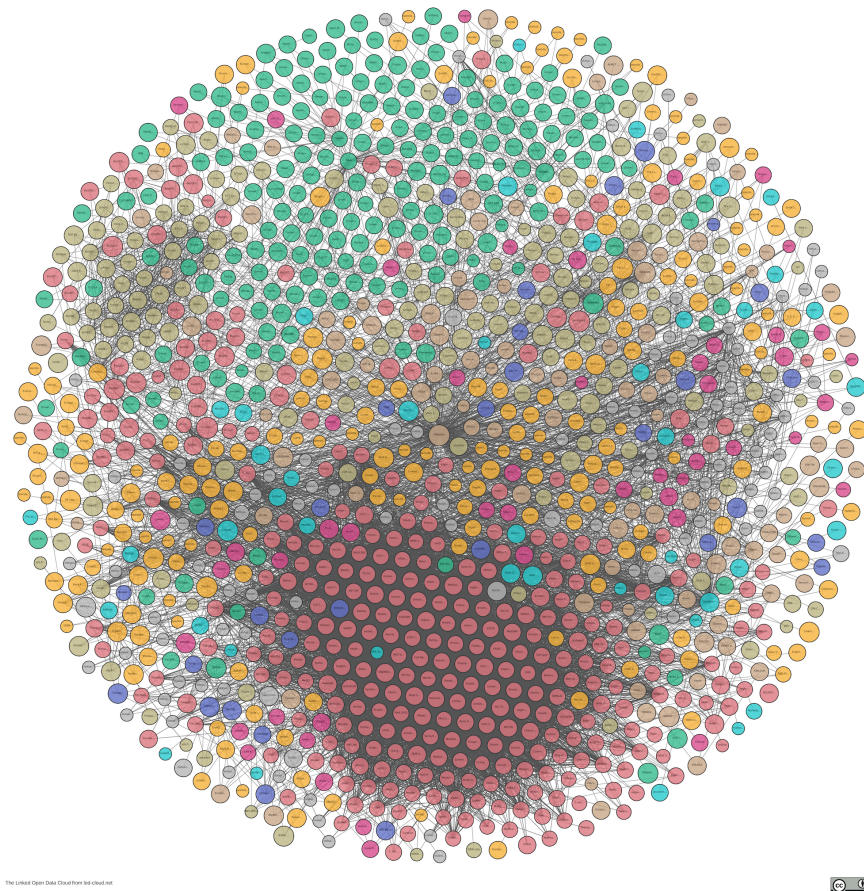
Providers are the first and recommended line of defense to educate the consumers about the HPV vaccine. Yet this is an added burden to the provider. One of the burdens is being a skilled communicator, especially toward patients who have low health literacy ([Goff et al., 2011](#); [Roter et al., 2007](#); [Evans & Bostrom, 2002](#)). Another burden is taking the role of being health educator ([Zimet et al., 2013](#)), even in the difficult situation of talking about sexual-related matters with both the parent and the teenage child ([Esposito et al., 2007](#); [Humiston et al., 2009](#); [Daley et al., 2006](#)). More importantly, there is not enough time to discuss and counsel the patient about the HPV vaccine ([Goff et al., 2011](#); [Vadaparampil et al., 2011](#)). This work explores the possibility of delegating and automating the communication and education task to an interactive kiosk or tablet application that is available for patients, similar to a “Siri”-like experience. The possible benefit other than cost effectiveness



is facilitating the interactive dialogue that could approximate a real conversation (Migneault et al., 2006), mimicking the personal interaction that patients and health consumers desire (Katz et al., 2011; Ahken et al., 2015). Also using machine intelligence, the system can strategize how and when information is provided (Migneault et al., 2006). For example, the system can implement the use of short sentences, acknowledge the user by name, supplement the speech with pictures and figures, or regulate the dialogue turn to five chunks of information at a time to accommodate cognitive limits. There is also the opportunity to reach out to multilingual patients or improve low health literate patients' confidence and communication (Migneault et al., 2006; Narayanan et al., 2004). Current research have noted that face-to-face counseling between patient and provider has an influence in accepting vaccination for the patient or the patient's parent (Kessels et al., 2012; Rambout et al., 2014; Reiter et al., 2009; Anhang Price et al., 2011). Research have also shown that interactive agents can improve educational gains of the learner compared to passive learning (Christel, 1994; Moreno et al., 2001; Moundridou & Virvou, 2002; Hongpaisanwiwat & Lewis, 2003; Guadagno et al., 2007; Van Mulken et al., 1999; Rickenberg & Reeves, 2000).

One of the originators of the web, Sir Tim Berners-Lee, envisioned a "web of data" in a seminal 2001 article for Scientific American (Berners-Lee et al., 2001). The central component of the semantic web architecture is the ontology layer. Ontologies are representational artifacts that link heterogeneous data, describe domain knowledge, and enable reasoning capabilities for machines to consume. From an artificial intelligence perspective, ontologies provide machines a vehicle to represent knowledge and share standardization of knowledge between agents. Biomedical researchers benefit from the tools and the inherit features that ontologies have to offer, ranging from the National Center of Biomedical Ontologies BioPortal, the Gene Ontology (Ashburner et al., 2000), etc. Many applications utilize biomedical ontologies for sophisticated data analysis. Yet in the age where copious health information is available on the web, there is no dedicated ontologies or formalized knowledge bases that target patients that can be used for knowledge acquisition or utilized for patient-centric

applications and tools. The development of the patient-centric vaccine knowledge base would be the first of consumer informatics-based ontologies, which could foreseeably be used in a variety of use-cases. Viewing the linked open data cloud (Figure 1.1), a web of connecting ontologies, most of the “professional level” ontologies predominate the health and life science ontologies.



**Figure 1.1:** Linked open data cloud. Licensed under CC BY 4.0. <https://lod-cloud.net/>

Also, Berners-Lee had presented his vision of an intelligent agent providing interactive, aggregate health data from the web to a casual users. While this vision has gotten closer to realization, we propose the possibility of an ontology-based agent that can fulfill a specific use-case, namely an interactive system using modern day speech technology to counsel patients on the HPV vaccine.

Such an agent, would utilize ontologies to coordinate the dialogue and visual interaction, and provide ontologies to query from for domain knowledge. In addition, I will also test the usability of the agent with prospective users to demonstrate some early perceptions and effectiveness. Using ontologies, there is a possibility to formalize knowledge, such as patient-level vaccine knowledge and standardizing various interaction modalities, that can be shared with other agents and be linked to other sources to expand the domain space.

CONTEXT IS KING    Permit me to have a candid moment. You, the reader, may be puzzled by the discussion of ontologies and why and how it would be used for HPV vaccine counseling \*. Even from my own subjective observation, the biomedical informatics community has a limited scope (or imagination) of how ontologies could be used in this area, much less be used in human-computer interaction. This is evident with the massive number of expressive reference ontologies that I mentioned. Research is so devoted to ontology-learning methods and knowledge engineering, but what about making these semantic technology usable, for what they were envisioned to do, than just being repositories of validated scientific knowledge?

I present the classic knowledge triangle in Figure 1.2. This triangle describes the evolution of information, from noise at the bottom to data (relevant) to information (processed, structured data), to knowledge (rules about the information), and to wisdom (apex of all knowledge). Modern technology like machine learning typically is used to classify information, occupying the information level of the triangle. Ontologies are associated with the knowledge level of the triangle. For further explanation, I present (again) the knowledge triangle (Figure 1.3), but oriented in a use-case that may be analogous to what we are attempting to achieve.

Let us pretend we have a robot in a concert arena. On the left end of the figure, the musician emits a lot of noise, some of which might be useful. If useful, the analog to digital converter transforms

---

\*This might make sense in the next few chapters.

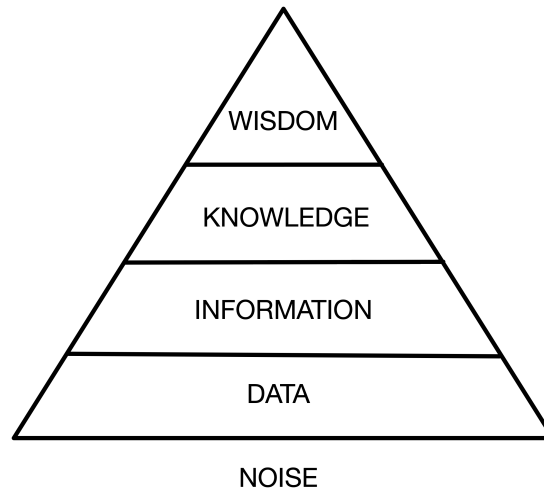


Figure 1.2: DIKW (Data, Information, Knowledge, Wisdom) pyramid

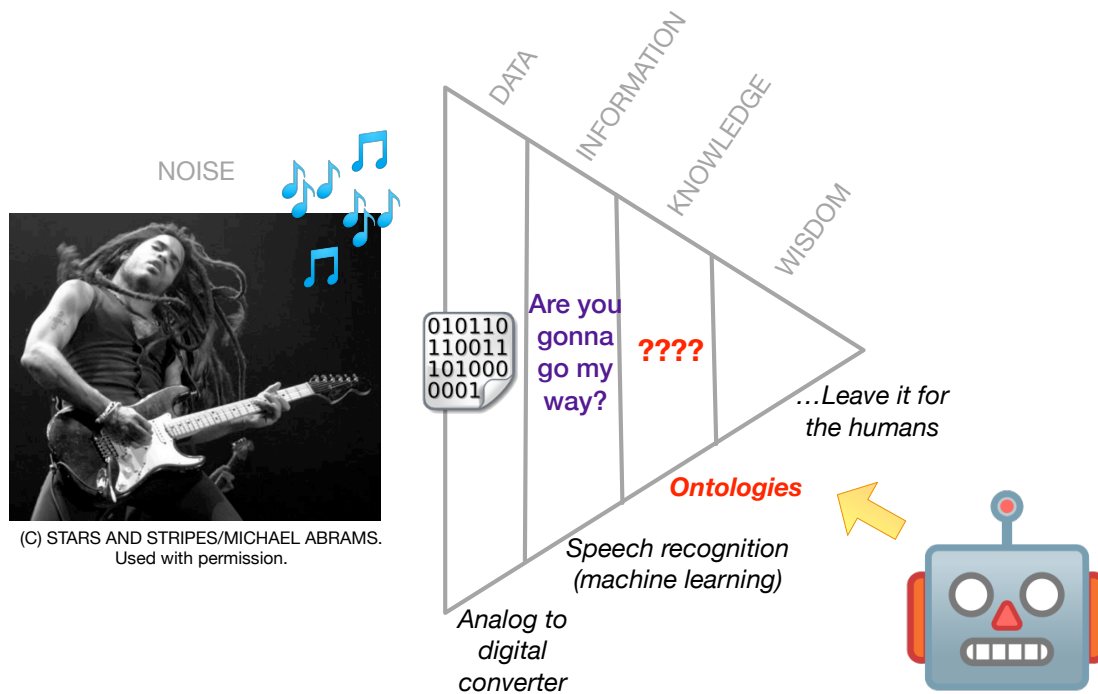


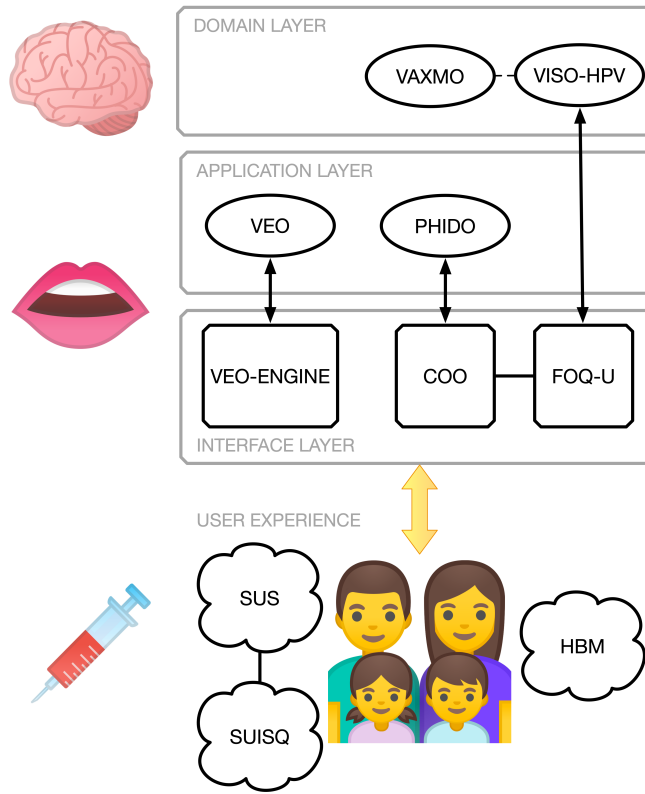
Figure 1.3: Application of knowledge pyramid for agents.

some of the noise to a digital format for the machine (data level). Machine learning and neural networks power the robot’s speech recognition software (information level) to find patterns from the digital data to produce string text, e.g. “Are you gonna go my way?” (Kravitz & Ross, 1993). At the knowledge level, given the string text of “Are you gonna go my way”, what does the robot do? Treat this as a question? Sing with the crowd? This is what ontologies for agents do, and why ontologies are important to not just our work but also for any system that requires rules for information and knowledge.

## 1.1 OBJECTIVE

Our overall objective is to certify that an ontology-driven architecture for the HPV vaccine impacts vaccination rates. By proving that this can work, we can (i) make a case to develop and deploy – based on the technology and lessons learned – a fully-automated conversational agent for the HPV vaccine in a clinical environment, (ii) further explore additional impact for vaccination uptake, and (iii) exploit the system for other health-related tasks.

The approach towards our objective is described in Figure 1.4. In that figure we have three integrated layers of the architecture for our system. The domain layer contains the domain knowledge base for the agent. The application layer contains application rules to automate the interface layer. The interface layer contains the basic functional units facing the user. This would include a dialogue manager to coordinate the counseling/conversation for the HPV vaccine, an accurate question answering system to address user questions about the HPV vaccine, and a visual facade for non-verbal behavior (emotion, regulating turn-taking, and feedback). This layer also harnesses the domain knowledge layer for some of its functionality. There is the user experience aspect that involve factors that could lead to vaccination uptake. While not technically part of the system architecture, the application and interface layer is inspired by early experimentation with a simulation agent. To



**Figure 1.4:** Multi-level development for the architecture of HPV vaccine conversational agent.

further support the objective, we also need to provide evidence that a conversational agent for the HPV vaccine is well received by the user, and that there is evidence to show that the agent may influence decisions for vaccine uptake.

## 1.2 RESEARCH HYPOTHESIS AND SPECIFIC AIMS

Our research asserts the following:

**HYPOTHESIS:** Application and domain ontologies can serve as the primary facilitator for the interaction of conversational agents for communicating and impacting vaccine knowledge in health

consumers.

**HYPOTHESIS:** A speech-enabled conversational agent can be an effective tool to enhance positive HPV vaccine beliefs.

#### **SPECIFIC AIMS**

1. **MODEL** patient-centric health knowledge relating to the HPV vaccine and serialize the structure of the knowledge into standard OWL and RDF format.
2. **DEVELOP** the interaction applications that harness application ontologies that can regulate the interaction flow between user and machine.
3. **ASSESS** the different tiers of the proposed conversational agent for HPV vaccine to validate its feasibility for field use.
  - Appraise the quality of health consumer ontologies for domain knowledge
  - Evaluate the system functionality of the ontology-driven application for dialogue interaction
  - Examine the overall system usability of the voice interaction interface through prospective users
  - Examine the impact of the conversational agent on the users' health beliefs of the HPV vaccine

### **1.3 DISSERTATION SUMMARY**

In Chapter 2, I review the background that supports this research. This chapter covers some of the basic history and the importance of the HPV vaccine on the human population and the need

to address the low vaccination rates, particularly provider interaction with patients and health consumers<sup>†</sup>. Chapter 2 also reviews the field of ontology and knowledge representation, along with the subfields pertaining to ontology evaluation and question answering for ontologies. Natural language processing research will also be briefly discussed as they are interdependent with ontology research. More importantly, dialogue systems are discussed from the perspective of utilizing ontologies, and from the perspective of dialogue systems impacting the healthcare domain.

Chapter 3 discusses development and design of health consumer ontologies for HPV vaccines. This includes the work that has resulted in the Vaccine Information Statement Ontology (VISO), Vaccine Information Statement Ontology for HPV (VISO-HPV), and the Vaccine Misinformation Ontology (VAXMO). This chapter will also explore my work in ontology engineering tools that involve using natural language generation and metrics underpinned by semiotic theory.

In Chapter 4, we review our work on applied ontologies for interaction with health consumers. This chapter highlights some preliminary work in implementing the Wizard of OZ protocol for data collection. Also, the chapter discuss the development and design of the Patient Health Information Dialogue Ontology (PHIDO), an application ontology for controlling machine decision for dialogue. Briefly, the Visualized Emotion Ontology (VEO) is mentioned in Chapter 4.

Chapter 5 embarks on assessing the different layers of the architecture of a conversational agent for HPV vaccine counseling. This chapter examines the evaluation of the first layer, namely the sufficiency of the domain ontologies – VISO-HPV and VAXMO. The chapter follows up on assessing the second layer which is the application tier. This involves reviewing the results of the evaluation of question answering software component (harnessing VISO-HPV) and the ontology-driven dialogue engine powered by PHIDO. Also discussed in this dissertation is the evaluation of PHIDO, and some brief notes about VEO and its own engine. The third layer pertains to the user experience

---

<sup>†</sup>Throughout this document, I will use patients and health consumers interchangeably as they are generally thought of as the same.



evaluation of the conversational agent where users interacted with the system and provided their feedback. The data collected is discussed in assessing the agent's usability and impact. Lastly, this chapter also explains the integration of the various tiers to serve as a blueprint for ontology-driven spoken dialogue system for HPV vaccine counseling.

Chapter 6 summarizes the contribution of this work, and its limitations and impact on future studies.

*Health informatics is the interdisciplinary study of the design, development, adoption, and application of IT-based innovations in healthcare services delivery, management, and planning.*

United States National Library of Medicine

# 2

## Background

This work covers two basic areas - domain and application ontologies and the impact of human papillomavirus (HPV) vaccines on health consumers. For the former, we will focus on the HPV vaccine, which until last year, was a vaccine that targeted adolescent boys and girls to protect them against the HPV virus. The vaccine was approved for females upon its inception in 2006 ([US Food and Drug Administration and others, 2006](#); [Garland et al., 2007](#)). In 2011, the Federal Drug Agency approve of its use for young males ([Castle & Scarinci, 2009](#)). Last year, the Federal Drug Agency approved the vaccine's effectiveness for up until the age of 45 ([US Food and Drug Administration, 2018](#)) (i.e. previously the vaccine had proven effectiveness for 11 to 26 years of age). Ontologies are electronic artifacts that codify the structure of knowledge to facilitate standardization of semantic knowledge and to provide reasoning intelligence for software agents. It is not uncommon to discuss

ontologies and natural language processing considering the use of semantic labels assigned to concepts and relationships. We will also discuss some subtopics of natural language processing that have a unique and important role with our research on health consumer ontologies, including dialogue systems. Overall, the topics in this chapter covers the background information in the later chapters of this dissertation.

## 2.1 HUMAN PAPILLOMAVIRUS VACCINE

At 79 million reported cases, human papillomavirus (HPV) infection is one of the most common sexually transmitted infections, and per year, according to the Centers for Disease and Control, 14 million more infections will be reported ([Centers for Disease Control and Prevention \(CDC\), 2013a; Weinstock et al., 2004](#)). Based on one report, HPV will infect 80 million women by the time they reach 50 years of age, and about 50% of sexually active males and females will eventually contract the HPV virus ([Centers for Disease Control and Prevention \(CDC\), 2011](#)). Also it is estimated that 80% of US women by the time they reach 50 will have been vulnerable to the HPV virus ([Chesson et al., 2014; Dunne et al., 2007; Weinstock et al., 2004; Centers for Disease Control and Prevention, 2016a; Chesson et al., 2014](#)). Previous figures noted that 20 million were infected by the HPV virus, and also projected, that 6 million more will be infected each year - a notable increase from what was described earlier ([Dunne et al., 2007; Weinstock et al., 2004; Centers for Disease Control and Prevention, 2016a; Dunne et al., 2007; Weinstock et al., 2004; Centers for Disease Control and Prevention, 2008](#)).

HPV virus sub-types are classified as low-risk or high risk, with the low-risk attributed to genital warts and the high risk attributed to cancer-related outcomes, such as cervical cancer, throat cancer, and genital-affected cancers ([Centers for Disease Control and Prevention, 2014b; Palefsky, 2010](#)). Specifically, the high-risk virus (HPV 16 and HPV 18) are accountable for 25-30% oral and throat

cancers, 90% of anal cancers, and 40% of penile cancers; and HPV 6 and HPV 11, the low-risk types, are primarily responsible for the occurrence of genital warts, which require comprehensive treatment (Parkin & Bray, 2006; Watson et al., 2008b; Lacey et al., 2006). Though uncommon, HPV infection can also be transmitted to children during pregnancy and delivery (Centers for Disease Control and Prevention, 2014b; Markowitz et al., 2014).

Reports have indicated an increase in cervical and genital cancer over the last decade, with 33,000 new cases of HPV-related cancer cases (Chaturvedi et al., 2011; Munoz et al., 2004; Centers for Disease Control and Prevention (CDC), 2012). Each year, 12,000 cervical cancer cases are reported in the United States, with 4,000 cases of deaths resulting from cervical cancer among women each year (U.S. Cancer Statistics Working Group, 2013; CervicalCancer.org, 2014; Centers for Disease Control and Prevention (CDC), 2012). Worldwide, the number of new cases of cervical cancer is estimated at 471,000, and it is the 5th most common type of cancer impacting females (CervicalCancer.org, 2014; Allen et al., 2011). While irrespective of the US region, women in the South are highly impacted by cervical cancer, and compared to other socioeconomic levels, women belonging to lower socioeconomic strata are greatly affected (Watson et al., 2008a; Hendry et al., 2013). Also, when compared to Caucasians, African Americans and Hispanics suffer more from cervical cancer (Reis et al., 2005; Newmann & Garner, 2005). In addition to cases of cervical cancers caused by HPV, a recent report estimates that 70-90% of new cases of head and neck cancer is due to the HPV virus (Young et al., 2015). While women can be diagnosed with head and neck cancer, non-smoking white males in their 40s and 50s are disproportionately affected by head and neck cancers caused by HPV (Deschler et al., 2014; Chaturvedi et al., 2011; Centers for Disease Control and Prevention (CDC), 2014; Hemminki et al., 2000). Regardless of demographic attributes, most of the population is susceptible to the HPV infection.

### 2.1.1 HPV VACCINE & POPULATION IMPACT

In 2006, the FDA approved of a HPV vaccine that can protect patients from low and high risk HPV viruses, specifically for patients for teens and young adults up to the age of 26 (US Food and Drug Administration and others, 2006; Garland et al., 2007). The HPV vaccine has shown its effectiveness in alleviating 70-75% of the occurrence of cervical cancer, while another study shows a success rate of 95% against cervical lesions associated with HPV viruses (Blake et al., 2015; Garland & Smith, 2010; Villa et al., 2007; Group et al., 2010). A few studies have also revealed efficacy of the HPV vaccine for women over 26, even though the HPV vaccine is licensed for an age group younger than 26 (Schwarz et al., 2011; Schiller et al., 2012; Castellsague et al., 2011). As of 2018, the HPV vaccine is approved for use up until age 45 (US Food and Drug Administration, 2018). Despite the success of the HPV vaccine, the United States is one of the “high income” countries to have low HPV vaccine coverage (28%) (Hopkins & Wood, 2013; Centers for Disease Control and Prevention (CDC), 2013c).

Irrespective of the efficacy of the HPV vaccine, it has yet to be fully embraced by the public, patients, and even healthcare providers (Gottlieb et al., 2014). The projected goal for Healthy People 2020 is 80% coverage for the HPV vaccine, but compared to other “high income” countries, the United States is falling behind these projections (28%)(US Department of Health and Human Services and Office of Disease Prevention and Health Promotion and others, 2012). Even with recommendations by the Advisory Committee on Immunization Practices (ACIP) to vaccinate adolescents, there are still a number of female teens and young adult females that are not completely immunized – either they have not had the vaccine, or not completed the three dose schedule (Markowitz et al., 2014). Infection rates for Black and Hispanic females are higher than females, resulting in greater incidence of cervical cancer when compared to Caucasian females (Hariri et al., 2011; Watson et al., 2008b). A few studies have pointed that non-white guardians and parents were less aware of the HPV infection

compared to white guardians (Constantine & Jerman, 2007; Hughes et al., 2009). One study in particular, noted that African American women were the least aware of the HPV virus (Cates et al., 2009). These disparities are further evident in vaccination rates compared to Caucasians, where white females have a higher vaccination coverage than African American and Hispanic. African American teens, in general, are also lagging behind Caucasian female teens and are the least likely to complete the 3 dose HPV vaccine routine (Rahman et al., 2014; Niccolai et al., 2011). Several studies have revealed that minority women, specifically black and Hispanic females were least likely to complete the HPV vaccine schedule (Rahman et al., 2014; Laz et al., 2013; Chou et al., 2011; Gelman et al., 2013).

Though not a representative study, one cohort group of males had 50% of the participants that have been infected by the HPV virus, while others vary around 20% - 73% infection rate among their samples (Giuliano et al., 2011; Dunne et al., 2006). Of concern, is that infected males tend to be asymptomatic and are expected to have more sexual partners than females (Dunne et al., 2006; Jones & Cook, 2008; Verhoeven et al., 2006). In one study, 30% of the infected males had the type of HPV virus that was attributed to cancer, while 38% had the genital related HPV virus (Giuliano et al., 2011). Currently, there is no FDA-approved test to detect HPV infection in males (National Cancer Institute, 2012).

Starting in 2009, the FDA approved the use of the HPV vaccine for males, yet current data revealed lowered coverage of males in the United States (Centers for Disease Control and Prevention (CDC), 2010; Jemal et al., 2013; Centers for Disease Control and Prevention, 2012; Reiter et al., 2011). The comparison with female teens (9-17) found that male teens had dramatically lowered HPV vaccination both for vaccine invitation and completion, 24% and 5% respectively in 5 US States (Du et al., 2015). While 2% of the boys, in another survey, had either completed or initiated the HPV vaccination (Reiter et al., 2011). Nationwide, the vaccine initiation and completion of the HPV vaccine with teenaged boys were 20% for initiation and 6% for completion (Rimer et al., 2014). Among the older

male demographic (19-26), the coverage compared to female is lower than females - 2.3% for males and 34.5% for females (Williams et al., 2014).

In one study, 80% of the teen male participants were unaware that the HPV vaccine was for males, but this lack of awareness is common with many other than teenage boys as Liddon, et al. noted in their review (Reiter et al., 2011; Liddon et al., 2010). However, a small percentage of parents (7%) either wanted to delay or reject the vaccine for their sons, but nearly all of them had not spoken to their sons about the HPV vaccine (Reiter et al., 2011). Contrasting with young adult males, a third or close to half of young males are open to the HPV vaccination, but some were undecided about the HPV vaccination (40%)(Ferris et al., 2008, 2009; Lenselink et al., 2008a). Liddon, et al., observed that the benefit of the HPV vaccine for the male's female partner does not encourage men for HPV vaccine uptake (Liddon et al., 2010). Health care providers and parents believe vaccinating males would be beneficial to female patients in alleviating HPV transmission, including avoiding stigmatizing females and advancing sexual responsibility for boys (Kahn et al., 2007; Olshen et al., 2005; Noakes et al., 2006; Katz et al., 2009). Overall, numerous papers have suggest wide support and advocacy of male HPV vaccination by parents and health care professionals (Kahn et al., 2009; Pearce et al., 2009; Songthap et al., 2009; Tariq et al., 2009; Katz et al., 2009; Noakes et al., 2006; Olshen et al., 2005; Wong, 2008; Dahlström et al., 2010; Lenselink et al., 2008b; Ogilvie et al., 2008; Podolsky et al., 2009).

### 2.1.2 CAUSES OF LOW HPV VACCINATION COVERAGE

Adult female patients have lower HPV vaccine coverage and are least likely to be vaccinated during adulthood than in their early teens (Centers for Disease Control and Prevention (CDC), 2013b; Markowitz et al., 2012; Goff et al., 2011). Specifically, studies show a range of 23-35% of adult aged women (19-26) have had one or two doses of the HPV vaccine, and 13% have completed the dosing interval (Williams et al., 2014; Laz et al., 2013). Researchers noted that college degrees, high income

earnings (>\$75000), routine health checkups, and perhaps regional locale are possible predictors of HPV vaccination completion (Rahman et al., 2014; Chao et al., 2010; Rahman et al., 2015; Du et al., 2015). Among women over 26, a survey indicates a majority of the women want physicians to discuss the HPV vaccine with an older demographic, and half want to be vaccinated with the HPV vaccine (Dempsey et al., 2015).

For patients, regardless of gender and age, factors that attribute to decisions to be vaccinated are perceived safety issues/side effects, recommendation from their provider, lack of vaccine knowledge, financial cost, and lack of vigilance (Williams et al., 2013; Kessels et al., 2012; Boehner et al., 2003; Ferris et al., 2009; Gerend & Barley, 2009; Ferris et al., 2008; Daley et al., 2010; Hernandez et al., 2010; Liddon et al., 2010; Reiter et al., 2010). Parents have a strong influence, particularly their personal perception and attitudes about the vaccine toward vaccine uptakes for their children, and for many parents and young adults, the lack of information was barrier to HPV vaccination uptake (Mullins et al., 2013; Rambout et al., 2014; Javanbakht et al., 2012; Perkins & Clark, 2012; Quinn et al., 2012; Burke et al., 2010; Dahlström et al., 2010; Fang et al., 2010; Podolsky et al., 2009; Rosenthal et al., 2008; Weisberg et al., 2009; Yeganeh et al., 2010; Kang & Moneyham, 2010).

A variety of studies have emphasized that a lack of information or knowledge gaps, regarding HPV and the HPV vaccine, is preventing parents with children, older patients, and children from wide regional surveys and demographics from receiving the vaccine, which ultimately may impact their overall health and lifestyle (Downs et al., 2008; Gottlieb et al., 2009; Dorell et al., 2011b,a; Laz et al., 2012; Stokley et al., 2011; Watkins et al., 2015; Etter et al., 2012; Fisher, 2012; Stupiansky et al., 2012; Priest et al., 2015; Marlow et al., 2007; Lenselink et al., 2008b; Woodhall et al., 2007; Al-Naggar et al., 2010; Bair et al., 2008; Bernard et al., 2011; Hilton & Smith, 2011; Krupp et al., 2010; Kwan et al., 2008; Toffolon-Weiss et al., 2008; Robbins et al., 2010; Liddon et al., 2012a; Sotiriadis et al., 2012; Juntasopeepun et al., 2012; Zimet et al., 2010; Mortensen, 2010; Kobetz et al., 2010a,b; Kontos et al., 2012; Rimer et al., 2014; Blake et al., 2015; Anhang Price et al., 2011; Dahlström et al., 2010; Williams



et al., 2011; Donadiki et al., 2014; Hendry et al., 2013). Interestingly, with patients who claim to have extensive knowledge of HPV infection, an examination of their knowledge reveals the opposite (Giede et al., 2010; Ahken et al., 2015). There are concerns among many parents and patients about the safety of the vaccine (Freed et al., 2010; Kester et al., 2013; Katz et al., 2009; Marlow et al., 2009c; Williams et al., 2011; Hopfer & Clippard, 2010; Madhivanan et al., 2009; Racktoo & Coverdale, 2009; Toffolon-Weiss et al., 2008; Brabin et al., 2006; Burke et al., 2010; Dahlström et al., 2010; Fang et al., 2010; Feemster et al., 2008; Ishibashi et al., 2008; McRee et al., 2010a; Riedesel et al., 2005; Rosenthal et al., 2008; Songthap et al., 2009; Wong, 2009; Yeganeh et al., 2010; Caskey et al., 2009; Marlow et al., 2009b). Additionally, the efficacy of the vaccine is a common worry among patients (Bakogianni et al., 2010; Liddon et al., 2012a; Hopfer & Clippard, 2010; Jain et al., 2009; Ratanasiripong, 2012; Juntasopeepun et al., 2012; Weiss et al., 2011; Zimet et al., 2010; Marlow, 2011). Globally, numerous publications have also noted that the false assumption that the onset of sexual activity upon vaccination was a concern for parents that may have resulted in hesitancy or refusal of the HPV vaccine (Humiston et al., 2009; Katz et al., 2009; Marlow et al., 2009c; Mays et al., 2004; Noakes et al., 2006; Olshen et al., 2005; Waller et al., 2006; Wong, 2008; Dinh et al., 2007; Tozzi et al., 2009; Askelson et al., 2010; Brabin et al., 2006; Marlow et al., 2009a, 2007; Brawner et al., 2012; Bastani et al., 2011; Hendry et al., 2013; Thompson et al., 2012; Brewer & Fazekas, 2007; Quintero Johnson et al., 2011). Because of this concern, health care providers are hesitant to discuss the HPV vaccine because of the implication of early teen sex. In one study, some health care providers are more comfortable discussing the sexual aspect of HPV and HPV vaccine with males than females (Schnatz et al., 2010; Zimet et al., 2013; Kahn et al., 2012; Ko et al., 2010). There is also a misconception among younger women and parents of daughters that if one is in a monogamous relationship, sexually inactive, or not engaging in risky sexual behavior, then the vaccine is not needed for them, and that they were at less risk than their counterparts (Zimet et al., 2010; Humiston et al., 2009; Kwan et al., 2008; Wong, 2008; Hilton & Smith, 2011; Hopfer & Clippard, 2010; Marlow et al.,

2009c; Katz et al., 2009; McClelland & Liamputtong, 2006).

For many patients and parents of patients, there is a perception of low susceptibility of HPV infection, and hence, they would not need the HPV vaccine. Some of these perceptions include being underage if the patient is a child, or that the HPV vaccine is not specifically needed for them (Wong & Sam, 2010; Bastani et al., 2011; Dempsey et al., 2009; Laz et al., 2012; Forster et al., 2010; Rambout et al., 2014; Robbins et al., 2010; Marlow et al., 2009c; Williams et al., 2011; Madden et al., 2012; Daley et al., 2010; Hernandez et al., 2010; Liddon et al., 2010; Reiter et al., 2010; Ferris et al., 2008). Instead, some patients revealed certain alternative measures, like condom use, cervical screening, or being in a monogamous relationship, precludes them from HPV vaccination (Hilton & Smith, 2011; Hopfer & Clippard, 2010; Mays et al., 2004; Noakes et al., 2006; Williams et al., 2011; Robbins et al., 2010; Henderson et al., 2011; Kwan et al., 2008; Leask et al., 2009; Waller et al., 2006; Caskey et al., 2009). It is also possible that HPV vaccination itself contributes to low susceptibility of cervical screening and protection against other sexually transmitted diseases (Henderson et al., 2011; Waller et al., 2006; Marlow et al., 2007; Robbins et al., 2010; Vamos et al., 2008).

Another factor contributing to low coverage of HPV vaccination is the lack of health care provider recommendation for the HPV vaccine. Often, if there is any provider recommendation, it is weaker compared to other vaccines that are recommended (Kester et al., 2013; Laz et al., 2012; Thompson et al., 2012; Stokley et al., 2011; Ylitalo et al., 2013; Centers for Disease Control and Prevention (CDC), 2013c; Vadaparampil et al., 2011; Grabel et al., 2013; Reiter et al., 2009; Zimet et al., 2010). One possible explanation is that physicians are unaware of the preventive impact that HPV vaccine has on cervical cancer and other cancers that result from the HPV virus (Saraiya et al., 2012; Koshiol et al., 2009; Perkins & Clark, 2012). Another explanation is the controversial issue of discussing the sexual activity aspect relating to the HPV vaccine, especially when the patient is a teenager (Esposito et al., 2007; Schnatz et al., 2010; Humiston et al., 2009; Kahn et al., 2007; Krupp et al., 2010; Sussman et al., 2007; Daley et al., 2006). Also, time constraints impede any in-depth or useful

discussion of the HPV vaccination (Vadaparampil et al., 2011). Similar to what patients assumed, some physicians believed that Pap tests are a preventive measure against cervical dysplasia, disregarding the threat the HPV virus poses (Perkins & Clark, 2012; Katz et al., 2011). Possibly due to language or cultural barriers, minority patients (African Americans and Hispanics) are least likely to receive HPV vaccine counseling (Jeudin et al., 2014; Centers for Disease Control and Prevention (CDC), 2013d; Hughes et al., 2009; Morales-Campos et al., 2013). However, with older patients, health care providers are likely to recommend the vaccine for adult females (Kahn et al., 2005; Riedesel et al., 2005; Daley et al., 2006; Kahn et al., 2005; Riedesel et al., 2005). A 2007 qualitative study noted that pediatricians believe there is greater benefit of the HPV vaccine for females than males, and that it also be difficult to recommend the vaccine for males (Kahn et al., 2007). Though irrespective of the health care provider recommendation, lack of access or regular checkup prevents patients from possible HPV counseling (Blake et al., 2015).

Several studies have mentioned cost of the vaccine from the patient perspective as a determinant to the HPV vaccine, even though the financial cost resulting from a cancer caused by HPV is much higher than the cost of being vaccinated (Pourat & Jones, 2012; Liddon et al., 2012a; Ratanasiripong, 2012; Juntasopeepun et al., 2012; Weiss et al., 2011; Zimet et al., 2010; Mortensen, 2010; Allen et al., 2009; Kim & Goldie, 2009; Wei et al., 2013; Williams et al., 2013; Tiro et al., 2012; Liddon et al., 2012b; Anhang Price et al., 2011). Notwithstanding high cost and possible insurance coverage, providers are still less inclined to provide the HPV vaccine (Pourat & Jones, 2012; Ylitalo et al., 2013). For half of the family physicians surveyed, the high cost of the HPV vaccine prevented purchasing them, and with some providers, issues with private insurance coverage deterred some from vaccinating insured patients (Freed et al., 2010; Colgrove et al., 2010; Vamos et al., 2008).

### 2.1.3 DIRECTIONS FOR INCREASED HPV VACCINATION UPTAKE

In 2014, the President's Cancer Panel suggested the need for patient-physician counseling for HPV vaccination, and for many patients, their health care provider is their main and most trusted source of health information to learn more about the HPV vaccine (Rimer et al., 2014; Hughes et al., 2009). Several studies have all declared that health care provider influence is an important factor too for HPV vaccine uptake for their patients (Kessels et al., 2012; Rambout et al., 2014; Reiter et al., 2009; Thomas et al., 2013; Lenehan et al., 2008; Anhang Price et al., 2011; Sundström et al., 2010; Williams et al., 2011; Rosenthal et al., 2011; Mullins et al., 2013; Reiter et al., 2009; Caskey et al., 2009). A few studies have noted as high as 95% vaccine acceptance whenever physician counseling occurs between patient and physician, and in one study an 18-fold probability increase of the acceptance as a result of health provider recommendation occurred (Hopfer & Clippard, 2010; Chow et al., 2010; Dinh et al., 2007; Fang et al., 2010; Kang & Moneyham, 2010; Caskey et al., 2009; Lau et al., 2012). Many patients prefer the face-to-face interaction and counseling to learn more about the HPV vaccine in order to decide on vaccine uptake (Katz et al., 2011; Ahken et al., 2015). While there is a strong preference for patient and physician interaction, this unfortunately forces the physician into taking on more intensive health education than time might allow (Zimet et al., 2013). This would involve additional professional development training and tips to address HPV vaccine barriers in order to effectively communicate health information to patients (Sherris et al., 2006; Valentino & Poronsky, 2015; Zimet, 2005; Kahn et al., 2007; Leddy et al., 2009; Fang et al., 2010; Mortensen, 2010). Also, this would require health care professionals to be aware of myths surrounding HPV vaccine, and detailing facts about HPV and the vaccine, and providing a comfortable atmosphere for patients (Valentino & Poronsky, 2015; Ahken et al., 2015). There is also the challenge that patients will ask few questions and interact minimally when being counseled on vaccines, as well as the health care provider dominating the discussion and peppering their dialogue with technical jargon that

could impede on the counseling (Goff et al., 2011; Roter et al., 2007). Yet, health care providers have limited time to discuss the HPV vaccine with patients, which also impacts decisions for vaccination uptake (Goff et al., 2011). In a study consisting of college males, participants noted that time and scheduling with their provider is a barrier that prevents opportunities to learn more about the HPV vaccine and HPV (Katz et al., 2011).

A couple of studies have noted that social interaction and support with peer and personal connection with physicians also has an influence on vaccine uptake (Allen et al., 2009; Hopfer & Clippard, 2010; Boehner et al., 2003; Ferris et al., 2009; Gerend & Barley, 2009). Policy or macro-level solutions, such as free vaccination of the HPV vaccine, does not seem to impact coverage of the population, as evident in Greece, where 11-25% of the population are vaccinated (Bakogianni et al., 2010,?; Donadiki et al., 2012). Mandatory vaccination policies could have a “knee-jerk” reaction in disrupting the trust between the public and its citizens (Gostin & DeAngelis, 2007; Charo, 2007).

To address lack of extended provider counseling and lack of knowledge of the HPV vaccine, educational intervention are needed for a wide range of patients who need to make a decision regarding HPV vaccination, especially those in remote rural areas where the resources to provide counseling are scarce, and minorities with lowered educational attainment (Blake et al., 2015; Herzog et al., 2010; Fiks et al., 2013; Watkins et al., 2015; Reiter et al., 2011). While the benefits of the HPV vaccine for their female partner does not encourage males to vaccinate themselves of the HPV vaccine, awareness and knowledge of how HPV infection and the vaccine impacts male health does (Jones & Cook, 2008; Boehner et al., 2003; Krawczyk et al., 2013; Gerend & Barley, 2009; Petrovic et al., 2011; Ferris et al., 2009; Lenselink et al., 2008a). One researcher noted that repeated exposure to HPV knowledge encouraged males to decide to receive the HPV vaccine (Daley et al., 2010). According to a review of HPV vaccination education initiatives, most of the educational materials are written documents that are distributed to patients (Fu et al., 2014). Most patients expressed discontent with the comprehension level and the quality of information in order to make an informed decision for HPV vaccine uptake

(Dempsey et al., 2009; Henderson et al., 2011; Madhivanan et al., 2009; Mays et al., 2004; Toffolon-Weiss et al., 2008; Williams et al., 2011; Wong, 2008). And overall, providing written documentation is insufficient to educate patients about HPV vaccination (Dempsey et al., 2006).

As far as the type of content that is needed to be communicated to patients, describing the HPV vaccine as a cancer preventive measure is a convincing argument for patients to eventually decide on uptake (McRee et al., 2010b; Krakow et al., 2015; Hilton & Smith, 2011; Hopfer & Clippard, 2010; Humiston et al., 2009; Madhivanan et al., 2009; Stretch et al., 2009). Also, for men, prevention of genital warts, protection for their partner and overall well-being as a result of being vaccinated provides forceful justification, along with the basic information about HPV infection, what the vaccine protects against, etc. (Ferris et al., 2008; Krawczyk et al., 2013). However, communication of HPV vaccination information whether verbally or written is complex and challenging, as one has to consider the information about sexual transmission (a sensitive topic for some patients), or communicating subtle information about the vaccine in a clinical environment (Quintero Johnson et al., 2011; Zimet et al., 2013). Too much information for the patient may pose some difficulties in a time constrained session or may cause fear in the patient, and caution should be considered when relaying negative information about the vaccine as it could impede on the patient's decision (Evans & Bostrom, 2002; Goff et al., 2011). In a survey among male college students, any information gathered about the HPV vaccine came from multimedia sources, even though participants preferred face-to-face counseling for the vaccine (Katz et al., 2011). Particularly, using web-based intervention for vaccine education and decision making has the benefit of being cost effective, flexible, interactive, and holds the potential to positively influence patient decisions regarding uptake (Lustria et al., 2007; Michael & Cheuvront, 1998; Brewer & Fazekas, 2007; McRee et al., 2010a; Fabry et al., 2011; Betsch et al., 2010). There has been no study to determine whether the web-based interventions and education ultimately led to actual vaccine uptake (McRee et al., 2012).

#### 2.1.4 HEALTH BELIEF MODEL

For any health intervention, grounding the intervention or program to any theory allows better design of programs and strategies to affect the behavior of the population and for improving and validating existing theories (Migneault et al., 2006). Particularly with interventions relating to either vaccines or vaccinations, theories of health behavior can potentially deduce beliefs that may lead to vaccine uptakes (Brewer & Fazekas, 2007). The health belief model (HBM) is one type of theoretical framework that aims to predict an individual's intent to engage in a preventive health behavior based on attitudinal and psychological factors. Simply, personally held beliefs can have an influence on whether one will undertake a particular preventive health behavior. HBM was introduced by Becker (1974) in the 1970s, and, as a theoretical psychological model, it has demonstrated predictive validity in various studies (Brewer & Fazekas, 2007; Zak-Place & Stern, 2004; Petosa & Jackson, 1991). Specifically with HPV vaccine research, studies have revealed that the health belief model has been shown to predict the intent to take the HPV vaccine among young females – for example Gerend & Shepherd (2012); Patel et al. (2012) and many other studies have utilized the health belief model to determine a subject's intent for vaccine uptake, specifically the HPV and influenza vaccines (Montgomery & Smith-Glasgow, 2012; Marlow et al., 2009b; Juntasopeepun et al., 2012; Coe et al., 2012). The variables that comprise the health belief model include perceived *susceptibility*, *severity*, *benefits*, and *barriers* (Becker, 1974; Janz & Becker, 1984).

- *perceived susceptibility* is “a person's subjective perception of the risk of contracting a particular condition or illness (perceived personal relevance)” (Bartholomew et al., 2011). If an individual believes he or she is of no risk of a harmful condition, they are least likely to behave in a manner to avert the condition.
- *perceived severity* is “a person's feelings concerning the seriousness of contracting an illness” (Bartholomew et al., 2011). A patient that assumes a condition is less severe is least likely to engage in healthy behaviors.

- *perceived benefits* is “a person’s belief regarding the effectiveness of various actions available to reduce the threat of a disease” (Bartholomew et al., 2011). A patient regarding a health behavior as a beneficial preventive measure against a condition or illness is likely to adopt the health behavior.
- *perceived barriers* is “a potential negative aspects of a particular health action” (Bartholomew et al., 2011). This includes factors, whether external or even psychological, that prevent an individuals’ intention towards a healthy behavior.

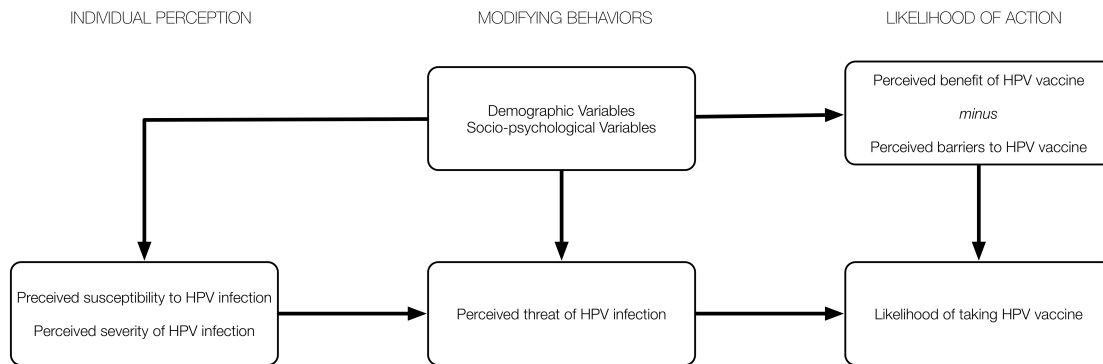


Figure 2.1: Adapted figure from Janz & Becker (1984) for HPV vaccine intervention.

In application to vaccination, these constructs translate to the risk associated of contracting a certain disease or infection if not vaccinated (*perceived susceptibility*), the assumption of the effectiveness of the vaccine (*perceived benefits*), the severity of disease or infection in absence of a vaccination (*perceived severity*), and the *perceived barriers* related to physical, or psychological barriers that would inhibit vaccination (financial or attitudes).

## 2.2 ONTOLOGIES

The word ontology has its roots in metaphysical philosophy, extending back to Aristotle’s *Categories*, as a “nature of being”. In the early 90s, the definition of ontology was oriented in the computer science field as a “specification of a conceptualization.” (Gruber, 1995). At the turn of the century,



Sir Tim Berners-Lee described his vision for the next generation web called the “semantic web” in Scientific America where ontologies would be the foundation for this vision (Berners-Lee et al., 2001). Simply, an ontology is a machine-readable artifact or file that encodes a logical representation of a domain space using vocabularies and their semantic meanings. It is the output of a knowledge engineering process where tools and methods are used to build the ontology (Gomez-Perez et al., 2006). Overall, ontologies are used for representing information and knowledge (Bodenreider & Stevens, 2006; Cimino & Zhu, 2006; Yu, 2006). Worth mentioning, ontologies still rely on their philosophical roots to help model the physical world as accurately possible in order for it to be used for science-related research, and there are currently two philosophical schools of thought, one of which is the realist approach, where the modeling of the physical world should be independent of our cognitive perception (Smith, 2003). The field of philosophy still plays an important role in ontologies as evident in the work of the Basic Formal Ontology (Smith & Grenon, 2019) which is an upper-level ontology that encodes philosophical concepts to align (and unite) the knowledge bases from biomedical ontologies.

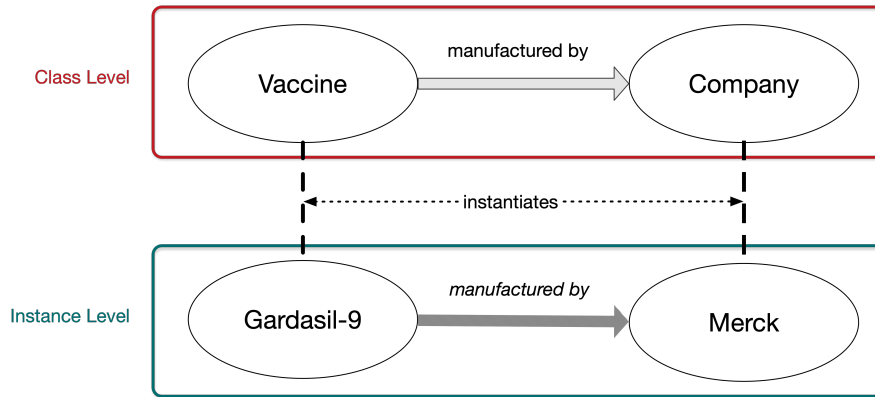
An ontology, according to Gruber, is a “description (like a formal specification of a program) of the concepts and relationships that can formally exist for an agent or a community of agents (Gruber, 1995).” Simply, an ontology consists of entities and relationships between the entities that define and model a knowledge space for a specific domain. Notationally, ontologies can be described in the following manner (Girardi, 2010):

$$O = \{C, H, I, R, P, A\} \quad (2.1)$$

Where  $O$ , in Equation 2.1 represents the ontology as a tuple.  $C$  is set of classes that abstracts the entities in the ontology, and  $I$  indicates set of instances in the ontology that instantiates the classes. Both  $C$  and  $I$  are generally entities of the ontology.  $H$  symbolizes the “*type of*” relationships or

hierarchical connections between classes ( $C$ ), and  $R$  represents the non-hierarchical relationships between the classes ( $C$ ) or instances ( $I$ ) - e.g. *protects*(*MMR\_Vaccine*, *measles*).  $P$  represents properties of the entities or their specific data types that they encapsulate, i.e. *has\_DOB*(*Patient*, *DateTime*), and  $A$  is a set of axioms that describe rules in the ontology, i.e. *Mother*  $\equiv$  *Parent*  $\sqcap$  *Female*.

A knowledge triple is a “building block” of an ontology that is comprised of two entities and at least one relationship between them - a *subject-predicate-object* statement. Figure 2.2 displays an example knowledge triple, showing the various dimensions of an ontology - class (TBox) and instance level (ABox). The class level from Figure 2.2 displays one knowledge triple *manufactured\_by*(*Vaccine*, *Company*) (*Vaccine is manufactured by a company*). On the instance level, *Gardasil-9* is an instantiation of *Vaccine*, and *Merck* is also an instantiation of *Company*. The relationship defined on the class level, *manufactured by*, is also alluded between *Gardasil-9* and *Merck*.



**Figure 2.2:** Example of ontology abstraction.

Not shown in the figure are assertions for relational knowledge, RBox. RBox are a finite set of object property assertions. These assertions provide additional meaning to the predication. For example, we can further define the *manufactured by* in Figure 2.2 by also defining another relationship, *manufactures*. We can define *manufactured by* and *manufactures* as an inverse property assertion. This would imply that not only a *Vaccine* is *manufactured by* a *Company*, but also asserts that a *Company manufactures*

a *Vaccine*. Thereby, the instance from Figure 2.2, could reason that Merck manufactures Gardasil-9. With our limited example, we can further express this relationship and the domain surrounding this triple with other RBox definitions:

- *SubObjectPropertyOf* - parent child relationship between classes and instances
- *EquivalentObjectProperties* - defines equivalent classes and instances
- *DisjointObjectProperties* - defines non-equivalency between classes and instances
- *InverseObjectProperties* - a relationship between classes and instances is an inverse of another relationship between classes and instances
- *ObjectPropertyDomain* - that a class or instance is domain of relationship
- *ObjectPropertyRange* - that a class or instance is a range of relationship
- *FunctionalObjectProperty* - a relationship that establishes that a class or instance (domain) has only one distinct class or instance (range)
- *InverseFunctionalObjectProperty* - inverse of FunctionalObjectProperty where a class or instance (range) has one distinct class or instance (domain)
- *ReflexiveObjectProperty* - defines a class or instance has a relationship to itself
- *IrreflexiveObjectProperty* - defines that a class or instance cannot have a relationship to itself
- *SymmetricObjectProperty* - a relationship between classes or instances is reciprocal
- *AsymmetricObjectProperty* - a relationship between class or instances in not reciprocal
- *TransitiveObjectProperty* - a relationship is between classes or instances that have indirect relationship with other related classes or instances.

Lastly, throughout this document, we use the word knowledge base and ontology interchangeable .

Formally, knowledge base is defined as only the TBox and ABox.

Being a tool to model complex information, specifically a domain like health care, yield some value for ontologies, along with querying complex information. Other benefits in using ontology to describe domain knowledge include (Yu, 2011):

- “Provid[ing] a way to reuse domain knowledge
- ...[M]akes domain assumptions explicit
- ...[P]rovides encoding of knowledge that machines can understand
- ...[A]utomat[e] large-scale machine processing...”

In general, knowledge in an ontology is represented as a triple in which information is presented as *subject*  $\rightarrow$  *predicate*  $\rightarrow$  *object*. Essentially, the *subject*  $\rightarrow$  *predicate*  $\rightarrow$  *object* are concepts that are “smallest, unambiguous unit of thought ... [that are] uniquely identifiable” (Groth et al., 2010). Each triple can seamlessly link to another triple to form an ontological knowledge base. For the knowledge to be readable by a machine, we use a computer-based syntax to encode it. Once encoded, this artifact can be shared and distributed for various purposes. Moreover, using OWL (Web Ontology Language) or RDF (Resource Description Framework), specific types of web ontology language syntax for ontologies, we can define more complex axioms and assertions to fully describe concepts which could provide machine reasoning capabilities.

The fundamental benefit of ontologies for the biomedical field or any other field is being able to connect research in biology with applications in medicine. Simply, by having machines and computational tools harness the encoded biomedical knowledge, it can reveal knowledge and information unrealized before, support the development of new software tools (Rubin et al., 2007), and model complex health information (Cimino, 1998).

One main benefit of an ontology is allowing for search and querying of aggregate heterogeneous data. Using a single identification where a concept might have several labels, heterogeneous information can be queried and retrieved if the information is associated with one of the labels. This is of particular use with image data where ontologies can link image data (e.g. MRIs) with non-image data (Martone et al., 2004), or of use with research data to map and link different data sources so they can be queried at once (Shah et al., 2006).

Another benefit is in information integration. Because ontologies can link and map different data sources, by nature, the information is integrated (Rubin et al., 2007). This benefit is of importance in the medical domain where there is a need to make data interoperable and shareable (Yu, 2006). Ontologies can also facilitate data exchange between systems. Ontologies are formalized representations of a domain space and are consistent formats that allow for distribution and sharing (Rubin et al., 2007). For clinical decision support systems or systems that utilize ontological knowledgebases, this would be a useful feature in sharing data between various systems (Yu, 2006).

### 2.2.1 ONTOLOGY EVALUATION

The evaluation of ontologies is not settled (Alani & Brewster, 2006; Obrst et al., 2007; Amith et al., 2018a). Ontology evaluation “is the problem of assessing a given ontology from the point of view of a particular criterion of application, typically in order to determine which of several ontologies would best suit a particular purpose” (Brank et al., 2005). For the last decade several ideas emerged addressing ontology evaluation (Brank et al., 2005), but none have appeared to be adopted universally by ontologists (Obrst et al., 2007; Almeida, 2009). Commonly, subject matter expert reviewers are sought to evaluate an ontology and provide feedback for changes. However, this effort is a time and resource intensive approach, especially if the reviewers need to acclimate themselves on the topic of ontology and ontology-related tools, like Protégé (Musen et al., 2015). A brief review of 200 randomly selected biomedical ontologies hosted on the National Center of Biomedical Ontologies’ (NCBO) BioPortal reveal that only 15 out of 200 have a formal assessment described in a corresponding design paper, and the remaining do not have any explicit documented evaluation (Amith et al., 2018a). With ontologies helping to further research in the biomedical domain, this highlights a strong need for evaluation for biomedical ontologies.

One of the future direction noted by Brank et al. (2005), is the need for automated tools to perform ontology evaluation. From research culled from the ACM and IEEE databases, six papers discussed

automated or semi-automated tools to perform ontology evaluation. All of the tools mentioned in the six papers were of experimental nature and not available for public use. Table 2.1 is from (Amith et al., 2019a) that briefly discusses some of the experimental tools for ontology evaluation.

Paper	Method
Ontology Evaluation and Ranking using OntoQA (Tartir & Arpinar, 2007)	OntoQA metrics (Tartir et al., 2005)
A Web-Based Ontology Evaluation System (Jianliang & Xiaowei, 2008)	Burton-Jones based; focused on the “subjective” metrics
A Survey on Ontology Evaluation Tools (Aruna et al., 2011)	Survey paper that discussed OntoAnalyser (OntoEdit plugin), OntoGenerator (OntoEdit plugin), WebODE plugin for OntoClean, Ontology Evaluation Tool, and S-OntoEval
Quality Model and Metrics of Ontology for Semantic Descriptions of Web Services (Zhu et al., 2017)	Paper is corrected version (Liu et al., 2016) that extends the ontology evaluation framework they introduced earlier
An Ontology Selection and Ranking System Based on the Analytic Hierarchy Process (Groza et al., 2014)	Applies analytic hierarchy process to evaluate ontology through Java-based application tools. Calculates language expressivity, domain coverage, size, consistency, and cohesion
Ranking ontologies in the Ontology Building Competition BOC 2014 (Jimborean & Groza, 2014)	Ranking-based metric system implemented as a web-based tool. Calculates structural, semantic, and term quality

**Table 2.1:** Papers surveyed for ontology evaluation software tools.

### 2.2.1.1 APPLYING SEMIOTICS THEORY FOR ONTOLOGY EVALUATION

Ontologies are sometimes alluded as symbolic representations of a domain space where the terms signify the entities contained within the domain space. Likewise, semiotics is a study of meaning behind signs and symbols or representations, divided by three aspects - *pragmatic*, *syntactic*, and *semantic*. Ontologies are generally thought of as semiotic products (Burton-Jones et al., 2005; Dividino

et al., 2008; Gangemi et al., 2006; Stamper et al., 2000) Rationale for utilizing semiotic theory have been explained in several practical endeavors. For example, Tolk et al. (2012) discussed the possibility of utilizing semiotic theory for modeling and simulation and Price & Shanks (2016) discussed the utility of semiotics for information systems and evaluating data models and information quality . Burton-Jones, et al. introduced an ontology evaluation framework based on the theories of semiotics that utilized various metrics formulated within the three branches of semiotics, along with an additional branch called “*social*” (Burton-Jones et al., 2005). Each evaluation criteria, based on the branches, asks if the ontology is useful (*pragmatic*), can it be read (*syntactic*), can it be understood (*semantic*), and can it be trusted (*social*). Each of these branches are decomposed to additional aspects that derive their values from data acquired from the ontology and external sources - subject matter feedback, number of links to the ontology, etc. The weighted sum ( $b_1, b_2, b_3, b_4$ ) for the four criteria results in the *overall quality* score (Equation 2.6).

$$P = b_{p1} \cdot PO + b_{p2} \cdot PU + b_{p3} \cdot PR \quad (2.2)$$

Equation 2.2 illustrates the calculation for the *pragmatic* quality ( $P$ ) where  $PO$  is the *comprehensive* quality value,  $PU$  is the *accuracy* value, and  $PR$  is the *relevance* value. The normalized sum of the three values provides *pragmatic* quality score. *Comprehensiveness* alludes to size of the ontology to cover a domain. The larger the ontology the more complete the ontology represents its domain. *Accuracy* is a rating for the truthfulness of the ontology, which is gathered from subject matter expert review, and *relevance* relates to compliance of user requirements.

$$S = b_{s1} \cdot SL + b_{s2} \cdot SR \quad (2.3)$$

The *syntactic* quality represented in Equation 2.3 takes the weighted sum of *lawfulness* ( $SL$ ) and *richness* ( $SR$ ). The *lawfulness* quality indicates compliance to the ontology language, and *richness*

value designates the amount of ontology features utilized.

$$E = b_{e1} \cdot EI + b_{e2} \cdot EC + b_{e3} \cdot EA \quad (2.4)$$

Equation 2.4 represents the *semantic* quality equation, where *EI* is *interpretability*, *EC* is *consistency*, and *EA* is *clarity*. The *clarity* value rates how ambiguous the terms in the ontology are, and *consistency* indicates consistent meaning of the terms. *Interpretability* measures term meanings.

$$O = b_{o1} \cdot OT + b_{o2} \cdot OH \quad (2.5)$$

The *social* quality score in Equation 2.5 is composed of the weighted sum of *authority* as *OT* and *history* as *OH*. *Authority* represents the number of ontologies that link back to the ontology, and *history* is the number of times the ontology has been accessed by the community.

$$Q = b_1 \cdot S + b_2 \cdot E + b_3 \cdot P + b_4 \cdot O \quad (2.6)$$

The overall quality score (Equation 2.6) is the weighted sum of *pragmatic* (*P*), *syntactic* (*S*), *semantic* (*E*), and *social* (*O*), which is a value between 0 (lowest) to 1 (highest). (Burton-Jones et al., 2005) contains details to derive the various values for each of the quality aspects to generate the *overall quality* score.

### 2.2.2 ONTOLOGY-BASED QUESTION-ANSWERING

Question answering (QA) is “the task of finding answers to natural language questions, meaning that question answering systems do not retrieve documents (like information retrieval systems), but instead provide short, relevant answers in an interactive setting” (Sonntag, 2010). Essentially, the aims of QA is to help users use natural language to find precise information and help end-



users query knowledge sources without having to code computer-level queries, using a natural language interface (Kaufmann & Bernstein, 2007). Some of earliest implementation of question answering were natural language interfaces for databases (NLIDB) (Androutsopoulos et al., 1995; Green Jr et al., 1961; Wood et al., 1978), and later, as data evolved into scheme-free corpora and natural language processing methods matured, question answering methods moved towards querying free-text sources and semi-structured sources (like health records) (Voorhees & Tice, 2000).

A completely different take on question answering is QA for ontologies. For the last decade, several QA for ontologies tools were proposed - AquaLog (Lopez et al., 2005), PANTO (Wang et al., 2007), NLPReduce (Kaufmann et al., 2007), Freya (Damjanovic et al., 2011), Querix (Kaufmann et al., 2006) - with relative success. These systems aimed to assist casual users and domain experts who have inexperience with writing query computer commands to easily query a knowledge information source, like a knowledge base triple store. While they each introduced their various approaches, they all exhibit some similar features. This included a gazetteer subsystem that build a list of terms utilized in the ontology, along with some procedures to perform term similarity between terms from the query and the gazetteer. Additionally, another similar feature among the QA systems is a process to extract knowledge triples from the natural language query, facilitated by a natural language parser or a combination of a few natural language methods.

There are benefits for an ontology-driven method for question answering over other question answering methods (Lopez et al., 2011). Compared to QA for databases, ontology-driven QA has the benefit of being loosely coupled from the natural language interfaces (NLI), meaning that there is minimal configuration work for the NLI to accommodate new knowledge sources or databases. Another benefit is semantics which allows for handling queries that have unfamiliar terms or terms that may have ambiguities, allowing for better precision in the responses for the questions. QA for free-text sources are burdened with taxonomic classifications for answer types, rule-based mechanisms to recognize types of entities, and use of information extraction to discern semantic relationships.

Whereas, in the ontology-based QA, these specific burdens are non-existent, and the ontology itself will have semantic relationships embedded. Other benefits also include (a) generating natural language answers from the triples, instead of reliance on snippets of texts that have to be ranked and chosen by a sub-system, (b) better handling unfamiliar answer types with the utilization of semantic relationships, and (c) the benefit of retrieving extended responses from other linked ontologies and/or inferred responses from reasoning capabilities of the ontology. Altogether, ontology-driven question answering has unique characteristics to improve on question answering research and a possibly provide better implementation.

### 2.3 NATURAL LANGUAGE PROCESSING

Earlier, the potential synergy between ontologies and natural language processing was noted. Natural language processing (NLP) is defined as “computational techniques for analyzing and representing naturally occurring languages at one or more levels of linguistic analysis for the purpose of achieving human-like language processing for a range of tasks or applications” (Liddy, 2001). The fundamental goal of NLP is to disambiguate and utilize natural language free text data for machine computational tasks. Within the context of this study, NLP research is supplementary to goals of this research. The body of research in NLP is expansive, ranging from question answering (discussed earlier within the domain of ontologies) to named entity recognition, sentiment analysis, etc. This section focuses on a few essential, yet unique, subtopics – one is the recent subfield of *open information extraction*, *natural language generation*, and primarily, *dialogue systems*. For the latter, we will focus on the spoken modality of dialogue systems, but we acknowledge other modalities like text-based (e.g. chatbots).

### 2.3.1 OPEN INFORMATION EXTRACTION

One aforementioned subfield relates to information extraction, which involves extracting predicates or triples automatically from unstructured sources for a variety of purposes such as ontology learning, text mining and question answering. Open information extraction (OIE) is one variation and is defined as “extraction paradigm that facilitates domain – independent discovery of relations extracted from text and readily scales to the diversity and size of the Web corpus. The sole input to an OIE system is a corpus, and its output is a set of extracted relations. An OIE system makes a single pass over its corpus guaranteeing scalability with the size of the corpus.” (Banko et al., 2007). It is generally domain independent, designed to handle a variety of unstructured sources –specifically content from the web that is heterogeneous in nature, and unsupervised (no training data required). Some open information extraction methods utilize dependency parsing that yields adequate precision and recall, while others utilize chunking, shallow parsing, and parts of speech tagging methods that yield high precision and low recall (Del Corro & Gemulla, 2013). Some examples of OIE methods introduced include TextRunner (Banko et al., 2007), NELL (Carlson et al., 2010b,a), ReVerb (Fader et al., 2011), OLLIE (Schmitz et al., 2012), CSD-IE (Bast & Hausmann, 2013), ClausIE (Del Corro & Gemulla, 2013), Stanford OpenIE (Angeli et al., 2015), Minie (Gashteovski et al., 2017), and Graphene (Cetto et al., 2018).

### 2.3.2 NATURAL LANGUAGE GENERATION

Natural language generation (NLG) aims to derive natural language text from data, and it is thought of as one of two main topics within the field of natural language processing (i.e. the other being natural language understanding (NLU)) (Reiter & Dale, 2000). NLG is defined as the “subfield of artificial intelligence and computational linguistics that focuses on computer systems that can *produce* understandable texts in English or other human languages” (Reiter & Dale, 2000). While

NLU focuses on derivation of structured data from free text, NLG focuses on producing free text from structured data. For the latter, the utility of NLG is to produce documents, ontology evaluation, and the summarization of data. Within the biomedical domain, when given large volumes of complex data, such as neonatal ICU data, NLG may be more beneficial than visualization of data (Portet et al., 2009). A variety of tasks or subtopics exist within the domain of natural language generation (e.g., structuring documents, content selection, lexicalization, co-reference expression generation, aggregation of sentences, and surface realization) (Perera & Nand, 2017). The main NLG task that our work focuses on is *surface realization*, specifically *grammar-based realization*. *Surface realization* is concerned with “mapping the text specification to the surface form of the sentences” (Perera & Nand, 2017) for human consumption. *Grammar-based realization* generally utilizes linguistic structures and heuristics to form sentences. The current state of the art in *grammar-based realization* is SimpleNLG (Gatt & Reiter, 2009) which offers an API interface to programmatically change the morphological and syntactical features for a generated sentence. Optionally, SimpleNLG can import the NIH Specialist Lexicon (Kazama et al., 2002) for biomedical related use-cases. We will refer to SimpleNLG in later chapters.

### 2.3.3 DIALOGUE SYSTEMS

At some point of the life of a normally developed individual, they will learn how to carry conversation beyond the obligatory “hello”. Determining whether to “give the floor” to another individual, figuring out what to say next, and how one should reply to someone’s speech are some of the basics of natural discourse. In speech, one can potentially communicate more information than in written form (D’Mello et al., 2010). Because speech is a natural act among humans, the ease to express thoughts in speech is relatively easier than written (Chafe, 1982; Tannen, 1982). In addition, speaking provides opportunity to convey more content in very little time (Damianos et al., 2003; Harris & Biermann, 2002; Litman et al., 2006). Machines, unlike humans, are not social entities, yet NLP

research is advancing the possibility of more interactive systems that can help users query systems. Yet managing dialogue using a computer-based system poses some challenges. [Jurafsky & Martin \(2009\)](#) states that dialogue is a form of discourse, which is a group of sentences that are connected by meaning and with entities that are clearly presented and described. Similarly, [Bickmore & Giorgino \(2006\)](#) also expound on the definition of discourse as “extended use of language to convey desires, beliefs, and intention” .

Within the scope of this proposal, a discussion of dialogue will be grounded in interactive learning between the educator agent and the learner. Dialogue can garner benefits by demanding the learner’s attention, assessing the learner’s gaps in knowledge and rectifying the gaps, and identifying miscommunication and incorrect information ([VanLehn et al., 2007](#)). Socratic teaching style, which in most forms involves interactive dialogue between student and teacher, is said to develop the student’s language skills ([Core et al., 2003](#)). Because speaking is a natural act, responses from learners require less effort than written format ([De La Paz & Graham, 1997](#)). One-on-one interactive learning also is said to provide better learning outcomes than formal classroom environments ([Cohen et al., 1982](#)), while not exercising pedagogical strategies ([D’Mello et al., 2010](#)).

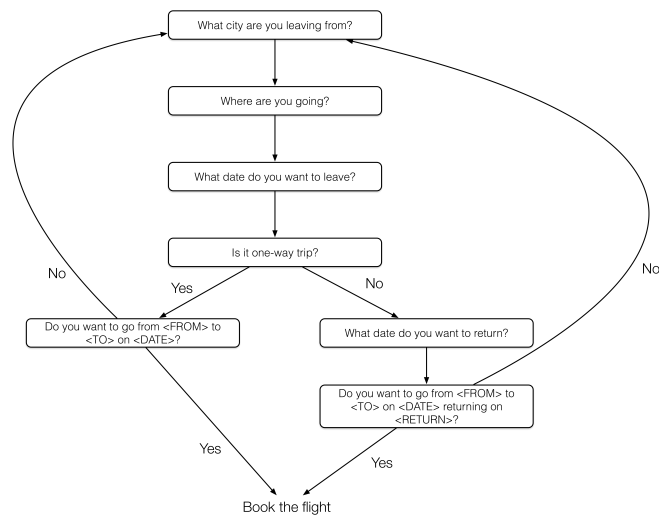
Regardless of whether it is computer-based or human-facing, interactive learning has demonstrated some educational gains with individuals ([Evens & Michael, 2006](#); [Graesser & Olde, 2003](#); [Lane & VanLehn, 2005](#); [Person et al., 2001](#); [Swanson, 1992](#); [Wood et al., 1978](#)). According to [VanLehn et al. \(2007\)](#), non-interactive learning is a state of just “pure reading or video watching without any problem solving or question answering”. Several studies, controlling for the same content in both the non-interactive learning group and the interactive learning group, revealed that larger learning gains are attained in the interactive learning intervention than the non-interactive group ([Evens & Michael, 2006](#); [Graesser & Olde, 2003](#); [Lane & VanLehn, 2005](#); [Person et al., 2001](#); [Swanson, 1992](#); [Wood et al., 1978](#); [Craig et al., 2004](#)).

To provide an interactive learning experience would require human resources which may be costly

(Allen et al., 2006). Dialogue systems can automate and mimic the face to face interactive experience for users, specifically for educational use cases. A dialogue system, based on the Journal of Dialogue Systems definition, is

“a computational device or agent that (a) engages in interaction with other human and/or computer participant(s); (b) uses human language in some form such as speech, text, or sign; and (c) typically engages in such interaction across multiple turns or sentences” (Konstantinova & Orasan, 2012).

Learning from an agent can be enjoyable and users may be more interested in learning from a computer agent (Johnson et al., 2000). Also of value, dialogue systems offer hands-free interaction, and the usage of expressive natural language (Allen et al., 2006). Dialogue systems allow very novice users to interact with a complex system or function using natural language (Allen et al., 2006). Despite users knowing they were interacting with a computer, they “felt” they were talking to a real “person”, according to interviews with users (Migneault et al., 2006).

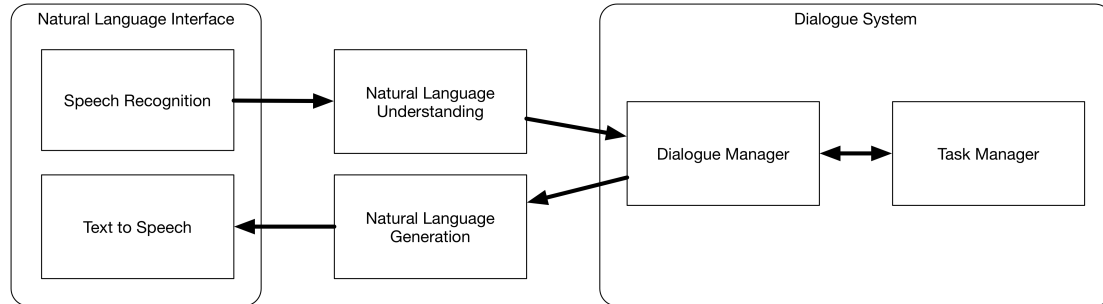


**Figure 2.3:** Simple implementation of dialogue management using finite state automaton (Jurafsky & Martin, 2009).

The literature describes at least 3 or 4 types of implementation of dialogue management - finite state, frame-based, and plan-based/inferential (Jurafsky & Martin, 2009; Ginzburg & Fernández, 2010; Bickmore & Giorgino, 2006). The finite state approach is the simplest approach where the nodes of a finite state automaton represent questions that the system would ask the user and multiple possible options and outcomes. Figure 2.3 taken from Jurafsky & Martin (2009) describes a simple architecture of the finite state based dialogue system. Unfortunately, finite state machine-based dialog management requires manually coding and thus is time consuming if modifications are needed (Kabanza et al., 2006; Rahati, 2012; Zhou et al., 1999). Another disadvantage is handling a large number of options or states in the dialogue and the designer has to anticipate every possible option in the dialogue discourse between the user and the system (Rahati, 2012). Frame-based dialogue management essentially captures missing information in order to execute a function. With a frame-based system, specific tasks require certain information captured from the user. Once, through discourse with the user, the system “fills” in the missing information, the system will execute a particular command. Plan-based/Inferential approach of dialogue management follows the same method as frame, but the salient difference is the use of inferences and assumptions. An advantage of using plan-based dialogue management is that it allows implementing discourse strategies and the use of templates to handle transitions as opposed to enumerating all options that finite state machines require.

Dialogue systems need to account for the speaker’s turn, classification of speech (speech acts), the subtleties of dialogue, grounding, and clear natural language speech. Figure 2.4 describes a basic architecture for a dialogue system courtesy of Jurafsky & Martin (2009). Natural language understanding requires the system to parse and analyze the natural language speech by the human participant, and natural language generation produces utterances in human understandable free text. Automatic speech recognition and text-to-speech synthesis are components that either translate utterances to an equivalent string of text (speech recognition) and text-to-speech synthesis takes natural language

free text and produces utterances from that free text. The dialogue manager and task manager are the components that coordinate the exchange between the system and user.



**Figure 2.4:** Basic architecture of a dialogue system (Jurafsky & Martin, 2009)

Another factor to consider, specific to health-related dialogue system, is to define a health objective and sub-objectives that lead to a health-related goal (Migneault et al., 2006). This could be a theoretical health intervention, like the health belief model (see 2.1.4).

Ontologies also offer some possibilities to extend the field of research for dialogue-based systems. One area is coordinating the sequence of discourse between the parties, where the ontology can index specific behavior in which the system can engage in (Johnson et al., 2000). Some of the other benefits inherent in ontologies, such as extensibility of knowledge space (Johnson et al., 2000) and providing reasoning and intelligence (Allnatt et al., 2001), can also contribute to dialogue management. In Section 2.4, we highlight the role of ontologies in agent-based systems.

#### 2.3.4 ONTOLOGY-BASED DIALOGUE SYSTEMS FOR HEALTH

In the health care domain, health dialogue systems have some prospective use cases to explore, which not only impact natural language processing research, but also medical and consumer informatics research (Bickmore & Giorgino, 2006). With a consumer's desire for face to face interaction and its overall effect on patient health (Naylor et al., 2004), health dialogue systems can support providers in areas where they are needed, specifically the lack of time providers have to interact with patients



(Davidoff, 1997), the lack of uniformity in communication style with patients, and providing detailed health information to patients.

Dialogue systems for health care imbues several benefits. In particular, health dialogue systems have the benefit of positively affecting the patient's health-related behavior and assisting in the observation of the health status of the patient. When offered as an alternative to paper-based documentation for patients, automated verbal communication can provide sophisticated goal oriented information delivery for patients (Bickmore & Giorgino, 2006). Health dialogue systems with the power of speech recognition can mimic the face-to-face interaction between provider and patient and automate that experience (Migneault et al., 2006). Specifically, the verbal mode of the health dialogue system can offer opportunities to enhance interactivity between patient and provider, such as using machine intelligence for decision making and coordination of content delivery, utilizing interpersonal cues to imitate human conversation and improve communication efforts with non-experts. It can also personalize the experience with the user (Migneault et al., 2006). Health dialogue systems can be cost effective if it is portable and generic, meaning if the system is not coupled with any specific domain (Allen et al., 2006). These systems also have the potential to reach a wider audience to deliver health information (Velicer et al., 1999). Communities that do not speak the native language of the provider can also be positively impacted by health dialogue systems that have multilingual support (Narayanan et al., 2004). In the field of serious gaming for health, games for health can also be enriched with dialogue (Luperfoy, 2004). While it may not replace the experience between patient and provider, it can help assist both parties - in helping the patient connect with other individuals and promoting self-management of care.

Since the 1990s, health dialogue systems, whether telephone-based or computer-based, have emerged in published medical research and demonstrated usage in a variety of health-based applications. Some examples of health dialogue systems utilization in managed care applications include nutrition (Farzanfar et al., 2004; Delichatsios et al., 2001; Glanz et al., 2003), cigarette use (Ramelson et al.,

2003), hypertension (Friedman et al., 1996), and asthma management (Adams et al., 2003). Also, health dialogue systems have been demonstrated in health behavioral interventions such as encouraging patients to engage in physical activity (Jarvis et al., 1997; Pinto et al., 2002; King et al., 2003), adhere to medication routines (Friedman et al., 1996; Young et al., 2001; Farzanfar et al., 2004), and encourage routine mammography screenings (Migneaul et al., 2005).

From a review of existing research for ontology-based dialogue systems on PubMed (Amith et al., viewb), only four papers involved the use of ontologies in some capacity. In one paper by Beveridge & Fox (2006), their dialogue system for breast cancer screening utilized an ontology as a domain knowledge base, while the dialogue flow is coordinated through an XML-driven system. Similarly, in another paper by Tielman et al. (2017), an ontology served as a knowledge base to link data collected from users suffering post-traumatic stress disorder through a virtual agent to build 3D worlds of their memory. The remaining papers, by Bickmore et al. (2013, 2011), utilized ontologies for conversational agents for diet interventions. Their work was somewhat limited as the ontology for speech tasks and the behavioral change theory ontology were not interoperable, and there was tightly coupling within their domain that prevented it from being portable.

From the body of literature discussed, it has been noted that provider-driven counseling on vaccines is an important task to support vaccine uptake. Research has indicated that these interactions do not take place for various reasons, even though consumers and other researchers find it effective. Dialogue systems have been experimented with since the 1990s and could offer a solution to automate the experience of conversation between machine and human. While it may not replace patient-provider communication, dialogue systems may offload the communication task and possibly help initiate a future discussion about the vaccine (or other health related topics) between the patient and provider in a subsequent meeting – a positive health literacy outcome.

## 2.4 AGENTS AND AGENT ARCHITECTURE

This subsection is light primer to software agents and their architecture towards describing and implementing them. Most of what is discussed here are relevant topics detailed by Wooldridge (2009). The term *agents* is an abstract definition of software system. According to Wooldridge & Jennings (1995), an agent “is a computer system that is situated in some environment, and that is capable of autonomous action in this environment in order to meet its delegated objectives”(Wooldridge & Jennings, 1995). Compared to other software systems, agents are, to a degree, aware of their surroundings incorporating sensory apparatus to retrieve feedback from environment where they inhabit. Based on the feedback the agent perform some action, and without any intervention from the builder or user, exhibits some autonomy.

Modern operating systems have software daemons which are basic examples of autonomous agents that operate in the background of every computer. More complex examples of agents are ones that exhibit some measure intelligence. To be considered as an intelligent agent, the system must have the following features (Wooldridge & Jennings, 1995):

- REACTIVE - respond to the environment in a timely manner
- PROACTIVE - initiative-driven, takes actions to fulfill objectives
- SOCIAL ABILITY - interacting with agents and humans to satisfy design objectives. This could involve negotiation and cooperation with autonomous agents with individual goals.

It is important to note that some software systems may be mistaken for being intelligent agents, like expert systems or even clinical decision support. Such systems are not agents due to being disembodied and are dependent on a user as an intermediary with the environment. For example, a social media

analytical tool powered by machine learning would not be considered an intelligent agent. Overall, agents act with relative autonomy and are not disembodied.

A variety of architectures have been proposed by researchers to design autonomous agents. Three basic agent architectures are *reasoning agents*, *reactive agents*, and *hybrid agents* which incorporates reasoning and reactive-based aspects.

**REASONING AGENTS** Reasoning-based agents are bifurcated as either *deductive reasoning* or *practical reasoning agents*. *Deductive reasoning agents* are theorem proving agents where it uses a database of domain predicates (the beliefs of the agent), and harnesses logical theory (the rules for the agent). They rely on representing the environment symbolically as logical formulae, e.g., *at\_location(desk)*. Overall they aim to solve :

- **TRANSDUCTION PROBLEM** - how to translate the world around agent into symbolic representation
- **REASONING AND REPRESENTATION PROBLEM** - applying (reasoning and manipulate) the symbolic representation for the agent

*Deductive agents* benefit from the clear, simple and transparent logical semantics. However, the drawback is the challenge of translating the attributes of the environment into symbolic representations. *Practical reasoning agents* utilize reasoning to preform an action by considering options and information than just theorem proving. In general, practical reasoning is driven to accomplish a goal through actions. These agents are involved in knowing the state of affairs of the environment (deliberation), then deciding the actions to achieve the state (means-end, i.e. planning). Deliberation aspect of practical reasoning agents models the options available and then filter for the appropriate options to fulfill the agent's intention. During deliberation, the agent factors in the beliefs, intentions, desires, and the current beliefs of the environment. Once options are provided, the system uses a planning

algorithm to choose the best option/action. One noted example is the Procedural Reasoning System which uses some pre-defined actions (plans) that define preconditions and goals of each plan.

**REACTIVE AGENTS:** Rodney Brooks, Ph.D., the progenitor of subsumption architecture for reactive agents, outlines his position for reactive agents ([Brooks, 1991a,b](#)):

- Intelligence is an emergent attribute
- Representation is not required to display intelligence in agents
- Reasoning is not required to display intelligence in agents

Reactive agents do not adhere to any use of symbolic reasoning approaches, but subscribes to the notion that intelligence emerges from the environment. In other words, intelligence is not embedded within the machine, like in a statistical model, but an expression of behavior with the environment.

Reactive agents suffer from engineering challenges (i.e. difficult to build), and the lack of representational model limits their scope to the local environment. Also, scalability is an issue as it is limited to 10 layers.

One of the most popular reactive agents architecture is Brooks' own contribution --- the subsumption architecture ([Brooks, 1986](#)). This architecture is a finite state machine arranged in a hierarchy. The architecture has sensors that triggers an action (layer). Each layer of the architecture encapsulates the lower layers. Essentially, a behavior or action (layer) includes, or subsumes, the behavior or action of the layer(s) that are beneath it. Because of the simple approach it can also be implemented on the hardware level if needed. Other reactive agent architectures include Pengi ([Agre & Chapman, 1987](#)), situated automata ([Kaelbling & Rosenschein, 1990](#)), and agent network structure ([Maes, 1991](#)).

**HYBRID AGENTS** Hybrid agents are the current architecture approach as of recent. Hybrid agents merges ideas from reasoning and reactive agents but implements them in layers. As layers, it is easier to see the intelligence represented, i.e. layers for reactive, proactive and social ability functions.

These layers are arranged in either vertically or horizontally, the former reduces complexity than the later. One example of hybrid agents is the 3T architecture that presents a layer for “reactive skills” (reactive agents) and deliberation (practical reasoning) and structured in vertical layers. InteRRap is another example of a vertical layered architecture (Müller, 1996). However, hybrid agents lack the semantic coherency that reasoning agents would offer.

#### 2.4.1 ONTOLOGIES FOR AGENT ARCHITECTURE

Wooldridge (2009) also mentions the internal data structures of the agent. The data structure’s role in the system is to provide the agent with decision making capabilities to perform autonomously in the environment. Specific roles for the structures may include representing the domain knowledge for the agent, providing information of the surrounding environment of the agent, and cataloging the previous actions of the agent (e.g. for the agent to learn).

According to Hadzic et al. (2009), these data structures could be manifested as a group of ontologies. Furthermore, they state 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).

In the upcoming chapters, we will detail the modeling and development of ontologies that provide the autonomous intelligence for our proposed software architecture of an ontology-driven conversational agent for HPV vaccine. The ontologies for the system will represent domain information for HPV vaccine knowledge (patient-level), and interaction information for the agent’s environmental and historical actions. This architecture will also include the software to execute these models using NLP methods to be discussed in separate chapters. Later on, we will demonstrate the practicality of the tool through preliminary simulation studies – which will also inform the design of our agent.

## 2.5 SUMMARY

This chapter explores the background literature for HPV vaccine and ontologies, along with related topics such as natural language processing. HPV vaccination rates are low among the population and sometimes lower among specific demographics like males, and young adults. Ontologies, representational artifacts that contain semantically encoded knowledge, can model complex information and share that knowledge base with other systems. Machines that employ ontologies can be imbued with reasoning power and provide a repository that can be queried. With this, ontologies can be extended in the area of dialogue management, an area that has seen minimal realization but has potential in aligning with health theory models. In Chapter 3, we will discuss our work to develop a domain knowledge layer that include (1) a vaccine ontology model from patient-level documents, (2) an ontology that represents vaccine misinformation and (3) some tools to assist in engineering ontologies.

*Human behavior flows from three main sources – desire,  
emotion, and knowledge.*

Plato

# 3

## Health consumer vaccine ontologies

THIS CHAPTER covers the core knowledge bases that form an information repository for the HPV vaccine conversational agent. Chapter 2 overviewed some of the legacy architectures for autonomous agents where many had a subsystem or layer within their design relating to a domain knowledge base. Within the proposed design of our conversational agent, the knowledge base is a set of ontologies that encode patient-level information about the HPV vaccine. In contrast to “professional ontologies”, our consumer-level ontologies may have information that are of concern for the health consumer ranging from basic vaccine information (age appropriateness for the vaccine, dosing requirements, etc.) to myths surrounding vaccines and the source of where those myths come from. Overall, consumer knowledge are on the level of patient-directed brochures or flyers that do not delve into



the in-depth science about vaccines or use the nomenclature that experts wield. From our own study where we analyzed vaccine-related corpora, the knowledge structures (i.e. mental models) for vaccines differ between experts and health consumers (Amith et al., 2017a, viewa). The development of two vaccine knowledge bases that we introduce in this chapter pertains to (a) general HPV vaccine knowledge that is covered at the time of their visit with their provider, and (b) modeling and cataloging vaccine misinformation. In a later chapter, we analyze the quality of these knowledge bases using the tools and methods that we discuss in this chapter.

### 3.1 ONTOLOGY EVALUATION TOOLS AND METHODS

Much of our work, and the foundation of this research, involved the use of ontologies to serve on various aspects of the architecture of the conversational agent for HPV vaccine. Aside from application ontologies that are measured by their ability to preform a task, domain reference ontologies need to be evaluated (verified and validated) (Gómez-Pérez, 2004; Vrandečić, 2009). Whether (semi-)automating the development of ontologies using ontology learning methods or manually curating them, the encoded information (intrinsic and extrinsic dimensions) has to be examined to ensure proper future utility of the ontology. Overall, ontology evaluation is an important facet of knowledge engineering as it supports their development and maintenance (Amith et al., 2018a). Chapter 2 discussed a known semiotic metric suite by Burton-Jones and colleagues that evaluates an ontology on various perspectives based on semiotic theory - a study of meaning behind signs and symbols. The usefulness of this metric suite is that it is not specified to a certain domain, it provides a “holistic” evaluation of the ontology that covers the intrinsic and extrinsic aspects, and it is generally easy to compute the scoring (Burton-Jones et al., 2005). A later discussion in this chapter indicates some evidence of its usefulness to evaluate an ontology through the use of OntoKeeper (Amith et al., 2018b, 2019a), our automated tool that is based on the semiotic metric evaluation.

Before discussing the domain knowledge base in this chapter, we summarize our venture into ontology evaluation that includes developing natural language generation software, creating a method for evaluating ontologies using the semiotic metric suite by Burton-Jones and associates, and the development of a comprehensive web-based tool that automates the process of evaluating an ontology.

### 3.1.1 NATURAL LANGUAGE GENERATION WITH HOOTATION

In Chapter 2, we introduced natural language generation as one of two major sub-fields in natural language processing. As opposed to natural language understanding, natural language generation (NLG) produces free text from data. This has the benefit of making complex data simplified for consumption and also has the benefit of being useful for ontology evaluation (Amith et al., 2017b). For example, if given the Turtle syntax, the complex information embedded in the syntax would be transformed into a human-readable sentence (Listing 3.1). This sentence would be accessible for non-technical domain users who would usually have difficulty utilizing Protégé (Musen et al., 2015) or web ontology languages to examine the veracity of the ontology (Rector et al., 2004; Funk et al., 2007). One of the challenge for ontology evaluation was how do we loop domain experts with very little experience in using ontology authoring tools or knowledge of the semantic languages to evaluate the ontologies.

```
# http://informatics.mayo.edu/TE0.owl#TE0_0000048
TE0:TE0_0000048 rdf:type owl:Class ;
  rdfs:subClassOf TE0:TE0_0000084 ;
  rdfs:label "Saturday" ;
```

**Listing 3.1:** Sample syntax coding from Time Event Ontology in Turtle syntax

We developed Hootation, a software library API that could translate the machine level encoding to free text which could potentially make it easier for domain experts to examine the veracity of an ontology (Amith et al., 2017b). Hootation was developed in Java and derived some of the implementation

from existing semantic web tools developed by the Agile Knowledge Engineering and Semantic Web (AKSW) Research Group (Bühmann et al., 2015). Hootation utilizes both OWL-API (v4) (Horridge & Bechhofer, 2011) and SimpleNLG (Gatt & Reiter, 2009) to process an OWL-file ontology into natural language statements. We also extended support for additional axioms types beyond the initial 12 to 18 (Hootation now supports 25 axiom types).

We wanted to determine if the natural language statements from Hootation could accurately produce free text information that subject matter experts could understand. This would enable us to use the technology for the validation of ontologies and also allow us to integrate the technology into a framework like the semiotic metric suite for ontology evaluation.

We enlisted the assistance of evaluators who had familiarity with one of the following ontologies:

- PEOPLE ONTOLOGY (Schwarz & Kurfess, 2011) is an ontology representing family relationships and often used as a teaching example for students to learn of the various axioms in OWL. This small ontology has 13 classes, 8 properties, and 9 instance level data elements.
- TIME EVENT ONTOLOGY (School of Biomedical Informatics Ontology Research Group, The University of Texas Health Science Center of Houston, 2019) is an ontology derived from the Clinical Narrative Temporal Relation Ontology (CNTRO) (Tao et al., 2010) and it involves the representation of temporal entities and their semantic relationships between those entities. Time Event Ontology (TEO) has 156 classes, 51 properties, and 8 instances, and it aligns to an upper-level ontology - the Basic Formal Ontology (BFO) (Smith & Grenon, 2019).
- INFORMED CONSENT ONTOLOGY (ICO) (Lin et al., 2014) is an ontology representing informed consent documents and processes and related concepts concerning US Common Law and medical ethics. Similar to TEO, ICO utilizes the Basic Formal Ontology (BFO) and was built behind the principles supported by the Open Biomedical Ontology (OBO) Foundry (Smith et al., 2007). ICO has 375 classes and 86 properties.

The Hootation library imported the above-mentioned ontologies and exported natural language

statements for each of the axioms contained in the ontology, along with the axiom type and the logical axiom notation (See Table 3.1). Each evaluator was provided with the outputted data for ontologies that they were familiar with and provided a rating to gauge the clarity of the natural language statement (1-3, 1=clear, 3=not clear), fidelity to the encoded logical axiom (binary, 1=yes, 0=no), and fidelity to the domain that it expresses (1-3, 1=agreement, 3=disagreement). Table 3.2 aggregates the results for clarity, natural language fidelity to the axiom and the axiom fidelity to the domain.

**Table 3.1:** Sample output showing axiom data from ICO.

Axiom Type	Logical Axiom	Natural Language Equivalent
SubClassOf	ICO_0000062 $\sqsubseteq$ ICO_0000073	every human subject unable to give informed consent is a human subject

**Table 3.2:** Evaluators results from examining natural language output. \* $(\mu, \sigma)$  \*\* % of “yes”.

Ontology	Clarity*	NL Fidelity to Axiom**	Axiom Fidelity to Domain*
People	1.19 (0.42)	90%	1.01 (0.14)
Time Event	1.32 (0.63)	92%	1.13 (0.38)
Informed Consent	1.28 (0.58)	95%	1.36 (0.64)
	1.26	92%	1.17

From our assessment of Hootation, evaluators reported that the generated statements had high level clarity, and we noted that with a legacy ontology like the Informed Consent Ontology, that some of the encoded axioms ( $n \sim 41$ ) did not accurately represent the underlying domain knowledge. Overall, we concluded that Hootation has a strong usefulness in the engineering of ontological knowledge bases as it provided an easy review of the encoded knowledge (Amith et al., 2017b). In the next section relating to OntoKeeper, our automated tool for ontology evaluation, we incorporated the Hootation API library to generate the natural language statements for domain experts to produce

the *accuracy* score.

### 3.1.2 BENCHMARKING ONTOLOGIES WITH SEMIOTIC METRIC SUITE

The semiotic metric suite provides an overall score based on an aggregate of the weighted sub-scores (*semantic, social, pragmatic, and syntactic*). Interpreting scores for some of the aspects can reveal certain strengths and weakness of the ontology. For example, a low *richness* score would indicate the ontology's low of usage of ontology axiom features or a high *comprehensiveness* score would indicate high domain coverage. However, determining whether an ontology is a marked improvement (and thus a contribution) over a similar ontology, or how it compares with other ontologies of the same type could provide further interpretation of the scores. [Burton-Jones et al. \(2005\)](#) mentions comparing an ontology with a library of other ontologies to further gain insight of the overall quality of the ontology. We experimented with the idea of using a set of drug related ontologies and a sample set of National Center of Biomedical Ontologies (NCBO) BioPortal ontologies to generate scoring data that could serve as a benchmark to use the semiotic metric suite to evaluate biomedical ontologies, as well as our own ontologies.

The NCBO BioPortal website is a federally-funded repository for a wide-range biomedical ontologies with approximately with over 769 ontologies hosted on their portal . In September 2015, we sampled a set of biomedical ontologies that had the most visits, according to NCBO BioPortal data ([Amith et al., 2018a](#)). This sample included 66 ontologies that were the most popular of that month (Appendix A), and we individually went through each ontology and created a merged ontology file that included their imported ontologies. Using prototype software that we developed that was a precursor to OntoKeeper ([Amith & Tao, 2015](#)), we calculated their semiotic ontology evaluation scoring. Table 3.3 has the composite scores of mean, min, max and standard deviation of the 64 ontologies (2 were removed due to processing errors).

In a published study, we further explored benchmarking libraries of ontologies by tailoring the

**Table 3.3:** NCBO Sample Aggregate Scores

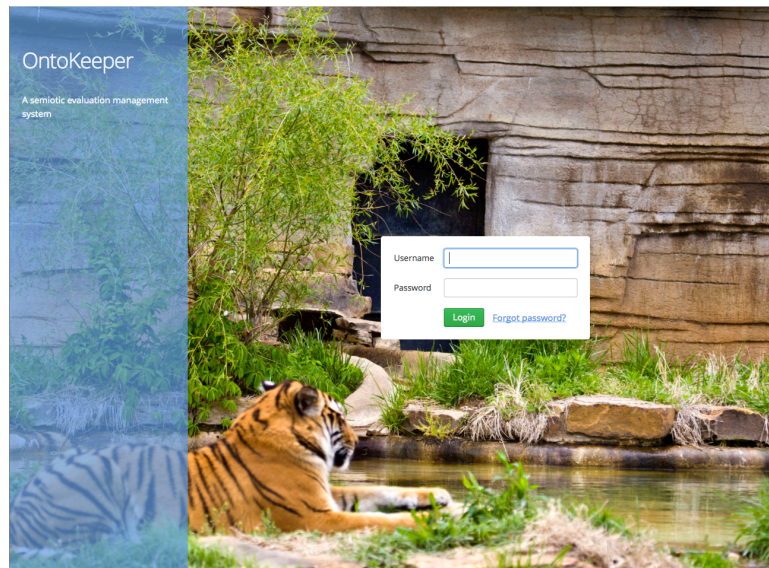
Quality	Mean	Std. Deviation	Min	Max
Syntactic	.64	.14	.18	.85
Lawfulness	.92	.16	.27	1
Richness	.36	.18	.07	.69
Semantic	.88	.15	.09	.99
Interpretability	.88	.14	.01	1
Consistency	.84	.40	-.17	1
Clarity	.96	.13	.14	1
Pragmatic	.02	.07	0	.52
Comprehensiveness	.02	.07	0	.52
Social	.02	.02	0	.13
History	.02	.02	0	.13
Overall Score	.39	.05	.21	.48

weights of the score to emphasize the strengths and de-emphasize the weakness of the ontology to attain a scoring that is representative of the ontologies (Amith & Tao, 2017). In that study we utilized a set of NCBO drug ontologies and produced an aggregate of semiotic scoring and produced a weighted formula that is specific for drug-based ontologies. This further demonstrates the value and flexibility of this metric suite to evaluate an ontology. Our work in ontology evaluation was advanced through the development of OntoKeeper which we utilized to evaluate the ontologies for the conversational agent, but also provide an on-demand evaluation of biomedical ontologies for other works (Manion, 2017; Lin et al., 2018).

### 3.1.3 ONTOKEEPER

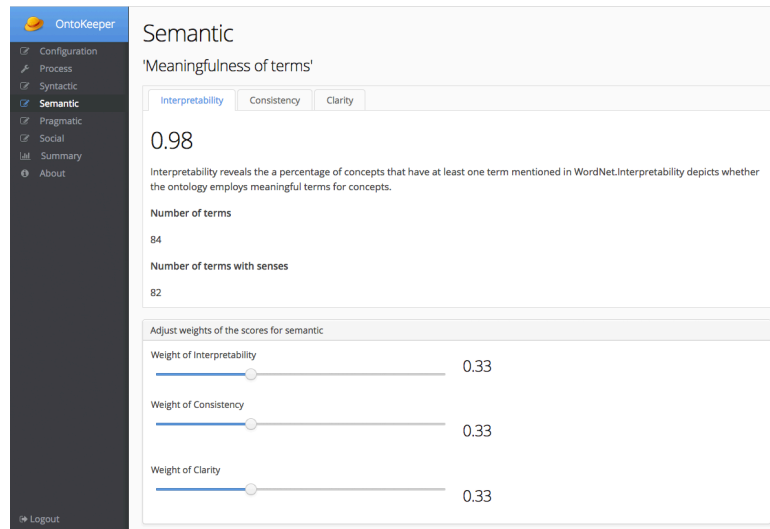
The previous works in ontology evaluation have culminated in the development of a web-based tool that automates the calculation of Burton-Jones et al. (2005) semiotic scoring metric for ontologies, called OntoKeeper (Amith et al., 2018b, 2019a) (See Figures 3.1, 3.2, 3.3). Chapter 2 noted the lack of automated tools to evaluate ontologies, and that future direction in this area of research would

entail developing the technology to perform this task (Brank et al., 2005). OntoKeeper grew out of a proof of concept prototype called SEMS (Semiotic Evaluation Management System) (Amith & Tao, 2015) and incorporated some of the lessons learned in previous research endeavors (Amith & Tao, 2017; Amith et al., 2017b; Manion, 2017).

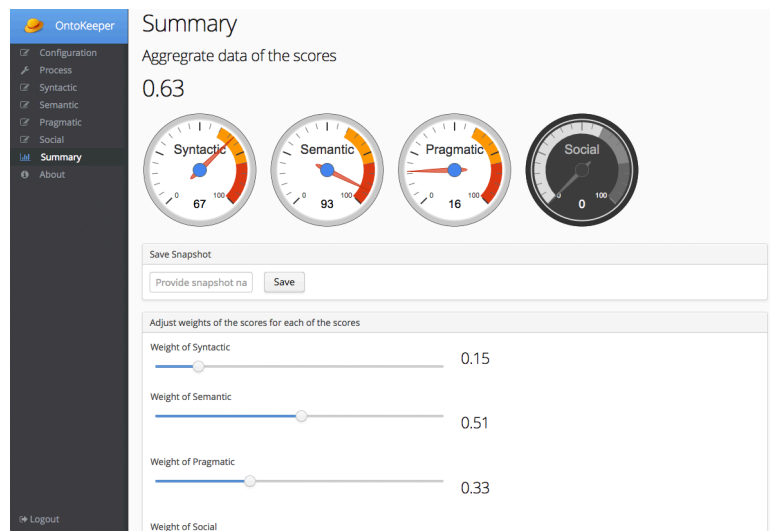


**Figure 3.1:** Login screen for OntoKeeper.

OntoKeeper is built on the Vaadin Java web framework (7.7) and incorporates various software libraries to handle the parsing of data from ontologies and processing of the labels - OWL-API (v5), MIT Java WordNet Interface (Finlayson, 2014), Apache Commons libraries, and Hootation. OntoKeeper also uses a PostgreSQL database server (v9.5.8) to store application data. The web application was primarily developed by myself and deployed for testing on an Eclipse Jetty web server hosted on an Ubuntu-based server (v16.04.3 LTS) with 4GB RAM with dual core processors. We enlisted five experienced ontologists with a background in engineering and publishing ontology artifacts. Each tested the tool using an ontology of their choice and then they all completed a System Usability Survey (SUS) (Brooke, 2013; Brooke et al., 1996), a simple, reliable usability survey that



**Figure 3.2:** One of four panels (Semantic score) to review the sub-scores for the semiotic metric suite.



**Figure 3.3:** Panel for the aggregated score

produces a score ranging from 0 to 100 (highest). The average scoring ( $\mu = 93.5$ ; minimum = 87.5, maximum = 100) from all of the participants equated to “A+” according to a percentile ranking (Sauro & Lewis, 2016).

Overall usability and the flexibility of the metric suite implemented through OntoKeeper provided



us with a tool to evaluate ontologies for this research (See Chapter 5). It also provides an automated tool needed for the biomedical ontology community to assess the quality of biomedical knowledge bases in the future.

In the next few sections, we discuss a couple of ontologies that provide the domain knowledge for the conversational agent for HPV vaccine - one relating to patient-level knowledge for the HPV vaccines and the other relating to representing vaccine misinformation.

### 3.2 DOMAIN KNOWLEDGE BASE FOR VACCINES

The subsequent sections cover the design and creation of two vaccine-related knowledge bases serialized as ontology artifact. The first ontology we discuss is the Vaccine Information Statement Ontology (VISO) and its derivative the Vaccine Information Statement Ontology for HPV Vaccine (VISO-HPV). The detailed effort in developing these ontologies have been documented in our previous studies ([Amith et al., 2015](#); [Wang et al., 2016](#)). Another ontology that is reviewed in the following section is the Vaccine Misinformation Ontology (VAXMO). Details about the development of VAXMO is also documented in our previous publication ([Amith & Tao, 2018](#)). The following sections summarize the inception and development of VISO, VISO-HPV (a contribution by Dennis Wang), and VAXMO.

The Vaccine Information Statement documents are federally mandated flyers given to patients at the time vaccination to inform them of the risks and benefits of a certain vaccine. They are “a document, produced by CDC, that informs vaccine recipients - or their parents or legal representatives - about the benefits and risks of a vaccine they are receiving” ([Centers for Disease Control and Prevention, 2014c](#)). Aside from informing the risks and benefits, it is also a minimal informed consent document to diminish any possible liability on the part of the health care provider. Through the National Childhood Vaccine Injury Act, these flyers must be provided to the individual getting the vaccine or

the parent/guardian of the individual ([Centers for Disease Control and Prevention, 2014c](#)).

The Vaccine Information statements themselves range from 1 to 2 pages ([Centers for Disease Control and Prevention, 2019](#)), and are written at 10th grade reading level ([Atkinson et al., 2012](#)). Currently there are 25 VIS documents available on the CDC website as PDF or rich text format, and they are provided in multiple languages including Spanish and Vietnamese. Also, we observed some general meta-level organization to these documents. Each VIS has a section for targeted population for the vaccine, probable reactions to the vaccine, dosing information and recommendation, and basic preventable disease information that the vaccine impacts.

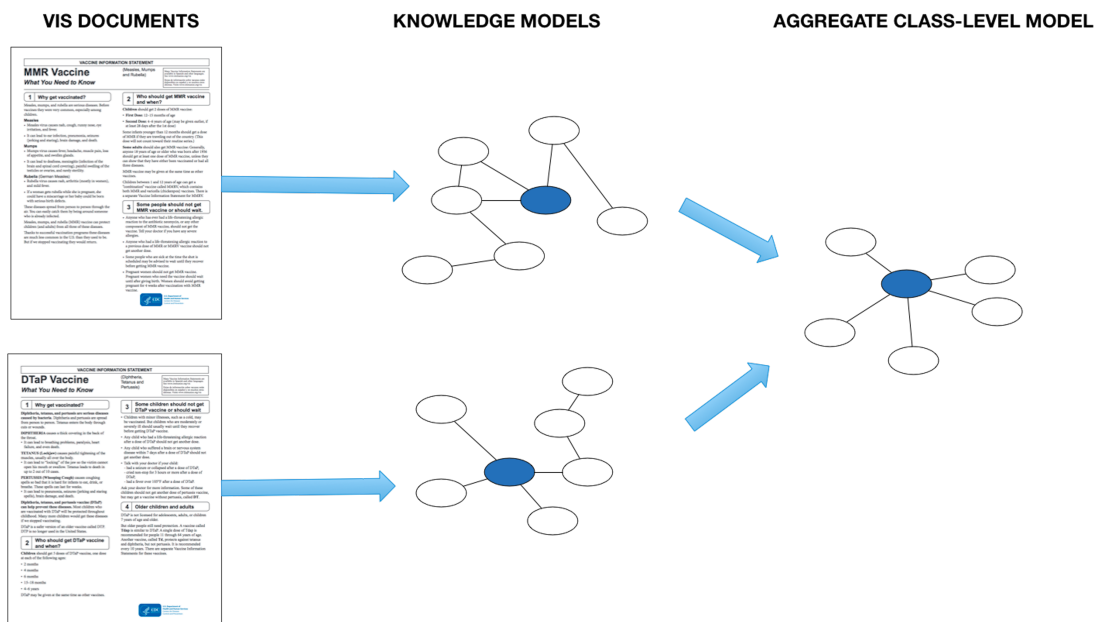
In context of this project, the goal of constructing the ontology of patient-level vaccine knowledge is to enable the conversational agent to provide foundational knowledge for questions about the HPV vaccine and possibly any other vaccine. By formalizing this level of knowledge we can represent the complexity of vaccine information that can be queried by machines and also provide a scalable information infrastructure to link to other knowledge bases. In a later chapter, we demonstrate the use of this ontology for the question-answering component. For conciseness, we discuss the TBox structure of VISO, but a detailed treatment of our work can be found in our previous literature ([Amith et al., 2015](#)).

### 3.2.1 VACCINE INFORMATION STATEMENT ONTOLOGY

We introduce the Vaccine Information Statement Ontology (VISO) that models knowledge contained in patient documentation for vaccines. The design of this ontology is based on analysis of six Vaccine Information Statement documents. These six VIS documents were:

- Diphtheria, tetanus, and acellular pertussis vaccine (DTaP) ([Centers for Disease Control and Prevention, 2016b](#))
- Rotavirus vaccine ([Centers for Disease Control and Prevention, 2016g](#))
- Hepatitis B vaccine ([Centers for Disease Control and Prevention, 2016d](#))

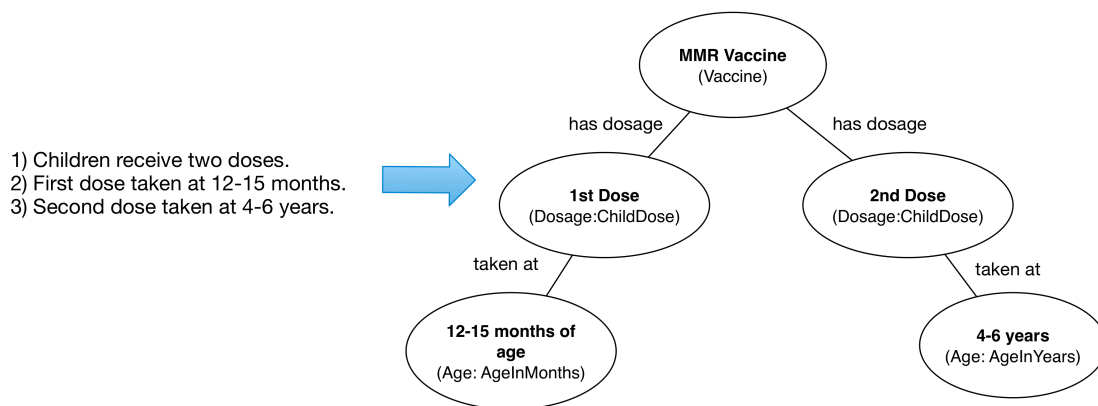
- *Haemophilus influenzae* type B vaccine (Hib) (Centers for Disease Control and Prevention, 2016c)
- Pneumococcal conjugate vaccine (PCV13) (Centers for Disease Control and Prevention, 2016f)
- Measles, mumps, and rubella (MMR) (Centers for Disease Control and Prevention, 2016e)



**Figure 3.4:** Process of creating models from each VIS document to create a singular common knowledge structure model.

The above-mentioned VIS documents were for the first six vaccines that any human should receive upon birth. The basic development task included extracting the concepts and entities, and constructing

the TBox level structure of the ontology. This would provide us with a framework to populate the ontology with instance data that machine can use to query or utilize for any task to be called upon. With each VIS document, we created a separate model from our analysis of the concepts, entities, and relationships identified. With each of these separate models, we aggregated the common high-level classes and relationships to create our TBox level ontology. The results of our analysis is shown in Appendix B in Table B.1 for the classes, and Table B.2 for the object property relationships. Table B.1 highlights the various class-level concepts contained in the Vaccine Information Statement Ontology, with 22 basic TBox-level concepts. We identified various subclasses to facilitate the categorization and descriptions from the VIS documents (e.g. VaccineAllergen and VaccineComponentAllergen for Allergen concept). Also we identified universal concepts that relate to the concepts from the VIS, like the Organ concept, which had several subclasses (e.g. Heart, Liver, Lung), and Substance (e.g. Gaseous, Liquid, Solid). We observed consistent use of scaling information, for example Mild, Moderate, and Severe (or Serious), to describe Reaction types and other concepts that utilized fuzzy labels. Further description of each of the concepts are detailed in the table.



**Figure 3.5:** Dosage Pattern from VISO.

From Table B.2, 33 property-based relationships hold between the concepts. We attempted to normalize relationship labels that were semantically similar within the domain of vaccine information, like

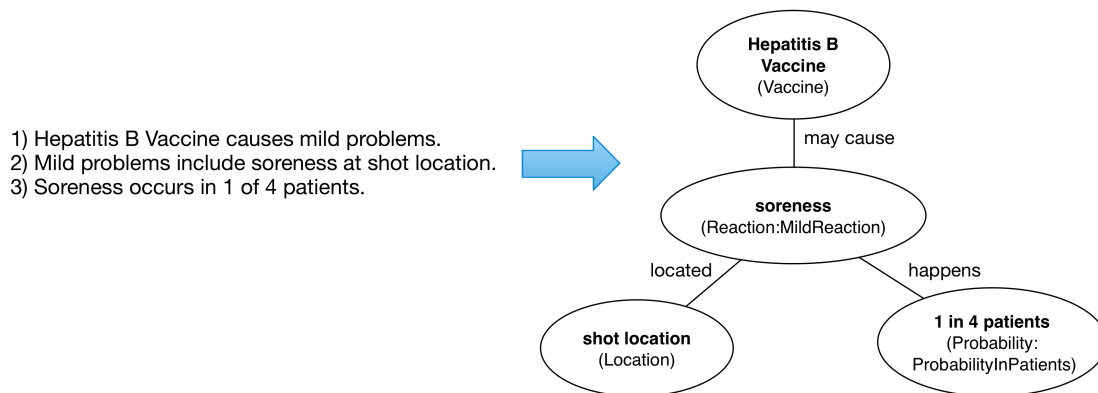


Figure 3.6: Reaction Pattern from VISO.

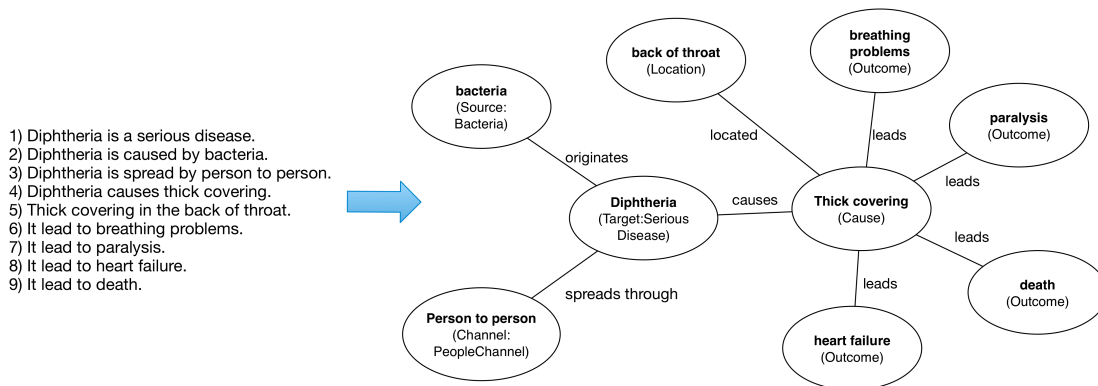


Figure 3.7: Target Pattern from VISO.

protects and prevents.

Aside from normalizing semantic labels for concepts and relationships, we attempted to structure consistent patterns that expressed the same knowledge. We noted three design patterns, one of them being for the vaccine dosing knowledge (Dosage Pattern, Figure 3.5). Another pattern for basic disease information was represented of how the targeted disease spreads, where it comes from, the cause, and what it can eventually lead to (Target Pattern, Figure 3.7). The third pattern was the Reaction pattern (Figure 3.6) that involves structuring knowledge about possible reactions, where it would be located, and the chances of attaining the reaction.

Figure B.1 in Appendix B shows simplified view of the TBox-level with the subclasses for readability purposes. Overall the TBox structure of VISO serves as the basic skeleton for instance-level data extracted from textual sources. From our analysis, we encoded the concepts and relationships in Protégé. This version of VISO is available at the SBMI Ontology Research Group website \*. The published version of VISO contained 132 classes, 33 object properties and 2 data properties.

### 3.2.1.1 VACCINE INFORMATION STATEMENT ONTOLOGY FOR HPV VACCINE

The Vaccine Information Statement Ontology for HPV Vaccine (VISO-HPV) is an extension of the Vaccine Information Statement Ontology (VISO). It uses the same TBox-level structure which has been upgraded since its inception. This version not only utilizes the same high-level structure, but it includes instance level data specific to the HPV vaccine from the latest Vaccine Information Statement document for the HPV vaccine and other sources. Detail of this work has already been published (Wang et al., 2016).

VISO-HPV has 38 instance-level data elements culled the VIS documents and Offit and Mosers's Vaccines and Your Child handbook (Offit & Moser, 2011). In addition, we have added a new annotation property called "natural language" which serves as a natural language translation of the predicate-level information. For example, the instance data of *cervical cancer*  $\rightarrow$  *death rate*  $\rightarrow$  *4000 women* had the natural language annotation of "*the death rate for cervical cancer is 4000 per year*". There was approximately 182 uses of the annotation in VISO-HPV. Figure 3.8 displays a screenshot of VISO-HPV in the Protégé editor.

---

\*<https://sbmi.uth.edu/ontology/project/viso.htm>

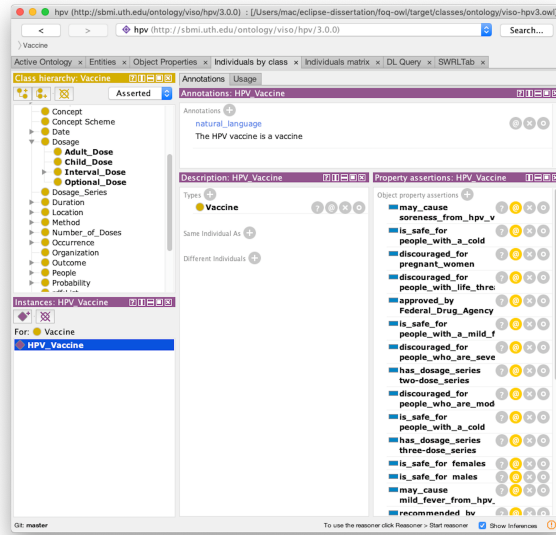


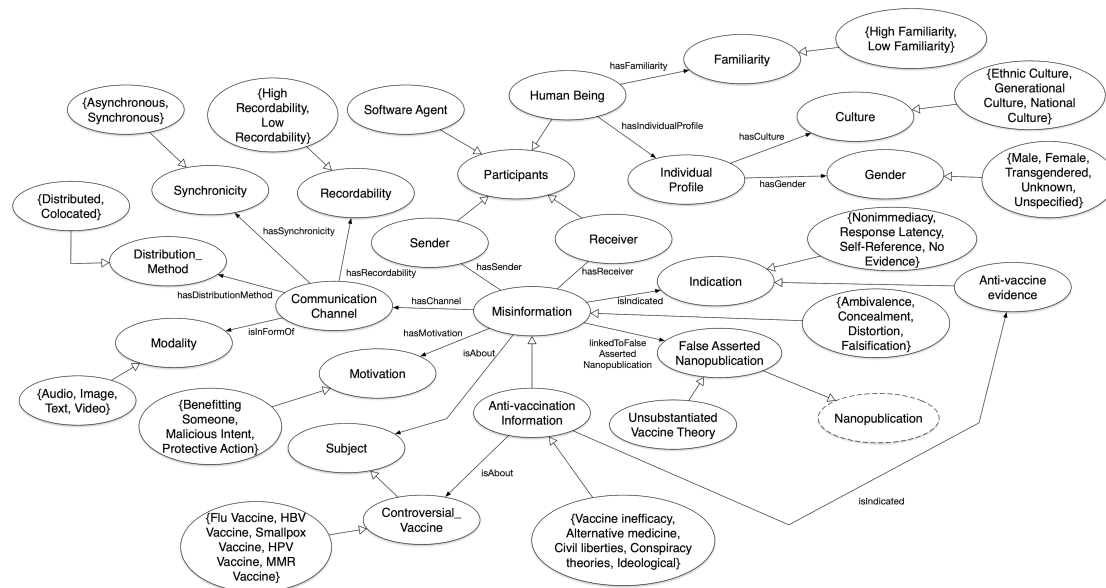
Figure 3.8: Screenshot VISO-HPV in Protégé.

### 3.2.2 VACCINE MISINFORMATION ONTOLOGY

First historical evidence of the root of vaccine misinformation may have emerged in the 19th century where the United Kingdom passed into law the Vaccination Act of 1953. As a result of the passage of this Act, a group (called the Anti-Compulsory Vaccination League) formed to retaliate and released publications that harbored anti-vaccination sentiments and beliefs. Years later, after vaccines have proven to be an effective measure against diseases, Andrew Wakefields' seminal but redacted paper alluding to the link between the MMR vaccine and autism had a major influence on vaccine discourse in the modern time (Opel et al., 2011; Burgess et al., 2006). Dr Wakefield's profile was also enhanced as he is still a major advocate for the anti-vaccine debate (Siegel, 2017).

Aside from public figures (Montanaro, 2017; White, 2015; Sears, 2011; InfoWars, 2017; Gorenstein, 2015; Greenfield, 2010; Kata, 2012), the Web has been a major tool for anti-vaccine advocates. In some early research by Kata (2010) and Bean (2011), early search hits when using certain vaccine-

related terms had anti-vaccination websites appear. In addition, social media has also been an influence on vaccination attitudes (Dunn et al., 2015; Dubé et al., 2016). Current news reports indicate pressure on tech companies to taper the misinformation content on their platforms but it is uncertain if it there will be an impact. Some surveys had some troubling figures, where 11% deemed vaccines as unnecessary and 20-25% believe in a link between autism and MMR Vaccine (Freed et al., 2010). To this day, measles outbreaks in various parts of the world, including North America, are not an uncommon news event (Lambert, 2019). One of the drivers for this project is to develop a knowledge



**Figure 3.9:** Class-level structure of the Vaccine Misinformation Ontology (VAXMO)

base of vaccine misinformation to equip the conversational agent with knowledge about pervading falsehoods about the HPV vaccine. Aside from that, the ontology can help machines understand





### 3.2.3 ANTI-VACCINATION INFORMATION CONCEPTS

Starting with the central concept of VAXMO, the Anti-vaccination Information concept, which is a subclass of the original Misinformation ontology (Figure 3.10). Since the Anti-vaccination Information concept is a subclass, it inherits all of the class definitions of Misinformation concept. The Anti-vaccination Information concept has several concepts like Conspiracy Theories, False Hoods, Alternative Medicine, etc.

```
# http://uth.tmc.edu/sbmi/vmo#Misinformation
<http://uth.tmc.edu/sbmi/vmo#Misinformation> rdf:type owl:Class ;
  rdfs:subClassOf [ rdf:type owl:Restriction ;
    owl:onProperty <http://uth.tmc.edu/sbmi/vmo#hasMotivation> ;
    owl:someValuesFrom <http://uth.tmc.edu/sbmi/vmo#Motivation>
  ] ,
  [ rdf:type owl:Restriction ;
    owl:onProperty <http://uth.tmc.edu/sbmi/vmo#hasReceiver> ;
    owl:someValuesFrom <http://uth.tmc.edu/sbmi/vmo#Receiver>
  ] ,
  [ rdf:type owl:Restriction ;
    owl:onProperty <http://uth.tmc.edu/sbmi/vmo#hasSender> ;
    owl:someValuesFrom <http://uth.tmc.edu/sbmi/vmo#Sender>
  ] ,
  [ rdf:type owl:Restriction ;
    owl:onProperty <http://uth.tmc.edu/sbmi/vmo#isAbout> ;
    owl:someValuesFrom <http://uth.tmc.edu/sbmi/vmo#Subject>
  ] ,
  [ rdf:type owl:Restriction ;
    owl:onProperty <http://uth.tmc.edu/sbmi/vmo#isIndicated> ;
    owl:someValuesFrom <http://uth.tmc.edu/sbmi/vmo#Indication>
  ] ,
  [ rdf:type owl:Restriction ;
    owl:onProperty <http://uth.tmc.edu/sbmi/vmo#linkedToFalseAssertedNanopublication> ;
    owl:someValuesFrom <http://uth.tmc.edu/sbmi/vmo#False_Asserted_Nanopublication>
  ] ,
  [ rdf:type owl:Restriction ;
    owl:onProperty <http://uth.tmc.edu/sbmi/vmo#hasEndTime> ;
    owl:someValuesFrom xsd:dateTime
  ] ,
```

```

[ rdf:type owl:Restriction ;
owl:onProperty <http://uth.tmc.edu/sbmi/vmo#hasStartTime> ;
owl:someValuesFrom xsd:dateTime
] ,
[ rdf:type owl:Restriction ;
owl:onProperty <http://uth.tmc.edu/sbmi/vmo#hasConfidence> ;
owl:qualifiedCardinality "1"^^xsd:nonNegativeInteger ;
owl:onDataRange xsd:double
] ;
dc:description "Central concept of the Misinformation Ontology by Zhou and Zhang" .

```

```

# http://uth.tmc.edu/sbmi/vmo#Anti-vaccination_Information
<http://uth.tmc.edu/sbmi/vmo#Anti-vaccination_Information> rdf:type owl:Class ;
rdfs:subClassOf <http://uth.tmc.edu/sbmi/vmo#Misinformation> ,
[ rdf:type owl:Restriction ;
owl:onProperty <http://uth.tmc.edu/sbmi/vmo#isAbout> ;
owl:allValuesFrom <http://uth.tmc.edu/sbmi/vmo#Controversial_Vaccine>
] ,
[ rdf:type owl:Restriction ;
owl:onProperty <http://uth.tmc.edu/sbmi/vmo#isIndicated> ;
owl:allValuesFrom <http://uth.tmc.edu/sbmi/ontology/vaxmo#Anti-vaccine_Evidence>
] ;
dc:description "Vaccine misinformation categories based on Kata's 2010 paper." .

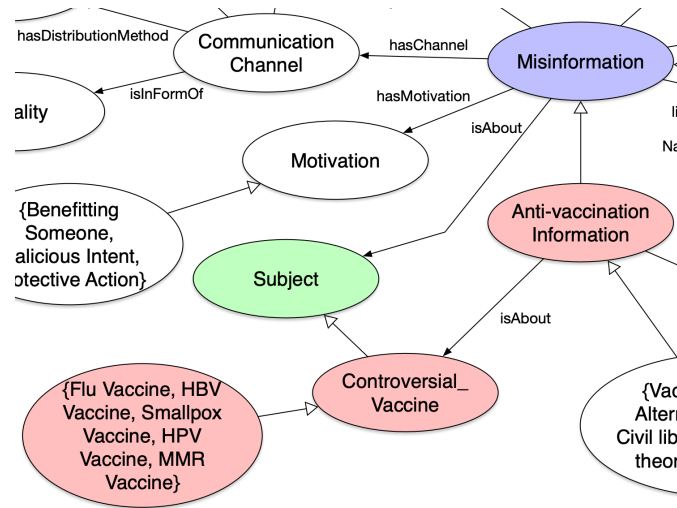
```

**Listing 3.2:** Turtle syntax of Misinformation and Anti-vaccine Information from VAXMO

### 3.2.4 CONTROVERSIAL VACCINE CONCEPTS

Figure 3.11 spotlights our modification to an appropriate concept label (green colored concept) of Subject, which was originally Object from the Misinformation Ontology. Subject is a concept that denotes what the misinformation is pertaining to, with a link (*isAbout*) between the Misinformation and Subject. As we noted earlier, Anti-vaccination information is a subclassed concept. Because of subclass inheritance, it has an analogous relationship with the Subject concept. In this case, Anti-vaccination information has association with Vaccine through the *isAbout* link to designate what vaccine the Anti-vaccination misinformation is concerning. The Vaccine concept has various

subtypes for various vaccine types, like HPV Vaccine, MMR Vaccine, etc.



**Figure 3.11:** Class structure of VAXMO showing the Subject concept and its associated classes, like Contraversial\_Vaccine. Listing 3.3 shows of the Turtle syntax representation.

```
# http://uth.tmc.edu/sbmi/mo#Subject
<http://uth.tmc.edu/sbmi/mo#Subject> rdf:type owl:Class ;
  rdfs:comment "Corresponds to Zhou and Zhang's object, \"object
    indicates what misinformation is false about\" (Zhou and Zhang,
    2007)" ,
    "relates to what the misinformation is about" .

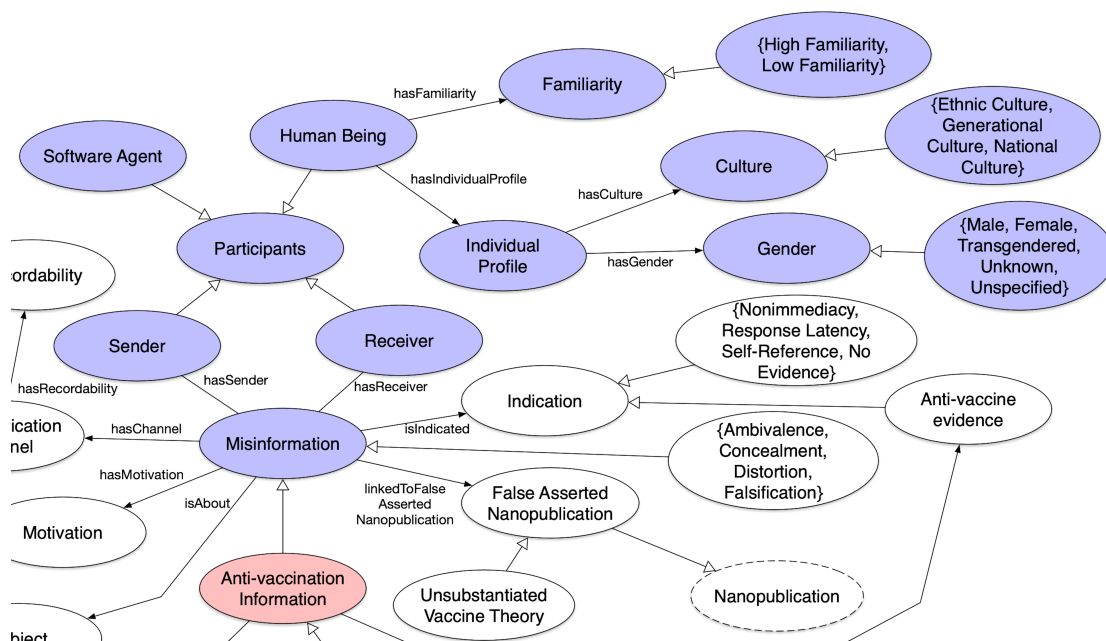
# http://uth.tmc.edu/sbmi/vaxmo#Controversial_Vaccine
:Controversial_Vaccine rdf:type owl:Class ;
  rdfs:subClassOf <http://uth.tmc.edu/sbmi/mo#Subject> ;
  dc:description "Limited focus on controversial vaccines that have
    caused contention among certain segments of the population." .
```

**Listing 3.3:** Turtle syntax of Misinformation and Anti-vaccine Information from VAXMO

### 3.2.5 INFORMATION THEORY CONCEPTS

Mentioned before, the Misinformation Ontology relies on concepts from the Information Theory process (Zhou & Zhang, 2007), like the inclusion of concepts such as Sender and Receiver. Figure

3.12 highlights some of the concepts from included the Misinformation Ontology. The Sender and Receiver classes were recreated and subclassed as Participants. Participants can also be a Software Agent, for example online bots and Human Being. The Human Being class is further extended with an Individual Profile concept that would define the concept culture, age, gender, etc. These concepts could be further elaborated with links to other ontologies that can define population information. The Familiarity concept (concept from Misinformation Ontology) for the Human Being concept describes how acquainted the individual(s) are with the misinformation.



**Figure 3.12:** Class structure of VAXMO showing concepts related to Information Theory process, which were derived from the original Misinformation Ontology. Listing 3.4 shows the Turtle syntax representation.

# <http://uth.tmc.edu/sbmi/mo#Participants>

```

<http://uth.tmc.edu/sbmi/mo#Participants> rdf:type owl:Class ;
  rdfs:subClassOf [ rdf:type owl:Restriction ;
    owl:onProperty <http://uth.tmc.edu/sbmi/mo#hasSize> ;
    owl:someValuesFrom xsd:int
  ] ;
  dc:description "agents involved in the communication of misinformation"
  ;
  rdfs:comment "originally MO involved just sender and reciever classes.
    According to Zhou and Zhang, \"Misinformation involves both the
    sender and the receiver, who are involved in producing and receiving
    misinformation, respectively. The sender of one misinformation
    instance may be the receiver of another one.\" .

# http://uth.tmc.edu/sbmi/mo#Software_Agent
<http://uth.tmc.edu/sbmi/mo#Software_Agent> rdf:type owl:Class ;
  rdfs:subClassOf <http://uth.tmc.edu/sbmi/mo#Participants> ;
  dc:description "A machine-based actor/participant who may be a sender or
    receiver of misinformation, e.g. chatbot, robot, etc." .

# http://uth.tmc.edu/sbmi/mo#Human_Being
<http://uth.tmc.edu/sbmi/mo#Human_Being> rdf:type owl:Class ;
  rdfs:subClassOf <http://uth.tmc.edu/sbmi/mo#Participants> ,
    [ rdf:type owl:Restriction ;
      owl:onProperty <http://uth.tmc.edu/sbmi/mo#hasFamiliarity> ;
      owl:qualifiedCardinality "1"^^xsd:nonNegativeInteger ;
      owl:onClass <http://uth.tmc.edu/sbmi/mo#Familiarity>
    ] ,
    [ rdf:type owl:Restriction ;
      owl:onProperty <http://uth.tmc.edu/sbmi/mo#hasIndividualProfile> ;
      owl:qualifiedCardinality "1"^^xsd:nonNegativeInteger ;
      owl:onClass <http://uth.tmc.edu/sbmi/mo#Individual_Profile>
    ] ;
  owl:disjointWith <http://uth.tmc.edu/sbmi/mo#Software_Agent> ;
  dc:description "homo sapien participant" .

# http://uth.tmc.edu/sbmi/mo#Receiver
<http://uth.tmc.edu/sbmi/mo#Receiver> rdf:type owl:Class ;
  rdfs:subClassOf <http://uth.tmc.edu/sbmi/mo#Participants> ;
  owl:disjointWith <http://uth.tmc.edu/sbmi/mo#Sender> ;
  dc:description "a participant who is on the receiving end of
    misinformation communication." .

```

```

# http://uth.tmc.edu/sbmi/mo#Sender
<http://uth.tmc.edu/sbmi/mo#Sender> rdf:type owl:Class ;
  rdfs:subClassOf <http://uth.tmc.edu/sbmi/mo#Participants> ;
  dc:description "a participant who is communicating misinformation" .

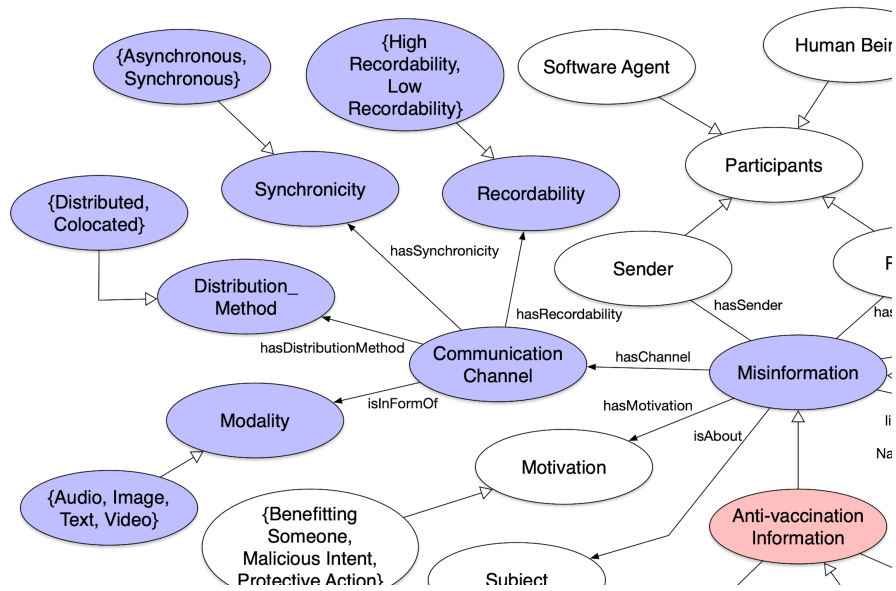
# http://uth.tmc.edu/sbmi/mo#Familiarity
<http://uth.tmc.edu/sbmi/mo#Familiarity> rdf:type owl:Class ;
  dc:description "\"The familiarity between a sender and a receiver of
  misinformation may influence the 'senders choice on what type of
  misinformation to produce and how to transmit it.\" (Zhou and Zhang,
  2007)" .

# http://uth.tmc.edu/sbmi/mo#Individual_Profile
<http://uth.tmc.edu/sbmi/mo#Individual_Profile> rdf:type owl:Class ;
  rdfs:subClassOf [ rdf:type owl:Restriction ;
    owl:onProperty <http://uth.tmc.edu/sbmi/vmo#hasCulture> ;
    owl:someValuesFrom :Culture
  ] ,
  [ rdf:type owl:Restriction ;
    owl:onProperty <http://uth.tmc.edu/sbmi/vmo#hasGender> ;
    owl:qualifiedCardinality "1"^^xsd:nonNegativeInteger ;
    owl:onClass :Gender
  ] ;
  dc:description "\"Prior deception research has suggested that some
  demographic features
  such as gender (Cody and Hair, 1983), age (Keating and Heltman, 1994),
  and cultural profile (Zhou and Lutterbie, 2005) of human information
  senders or receivers have impact on deception.\" (Zhou and Zhang,
  2007)" ;
  rdfs:comment "aka \"Persona\" in marketing" .

```

**Listing 3.4:** Turtle syntax of Participants and Familiarity concepts from VAXMO

Communication Channel (Figure 3.13) is also another legacy concept from the Misinformation Ontology. Information or misinformation is transmitted through channels. The Communication Channel describes how the information is exchanged and disseminated with concepts like Synchronicity, Modality, and Distribution Method. It is also entails the degree of accessibility through the Recordability concept



**Figure 3.13:** Class structure of VAXMO showing Communication Channel, a concept from the original Misinformation Ontology. Listing 3.5 shows of the Turtle syntax representation.

```
# http://uth.tmc.edu/sbmi/mo#Communication_Channel
<http://uth.tmc.edu/sbmi/mo#Communication_Channel> rdf:type owl:Class ;
  rdfs:subClassOf [ rdf:type owl:Restriction ;
    owl:onProperty <http://uth.tmc.edu/sbmi/mo#hasRecordability> ;
    owl:someValuesFrom <http://uth.tmc.edu/sbmi/mo#Recordability>
  ] ,
  [ rdf:type owl:Restriction ;
    owl:onProperty <http://uth.tmc.edu/sbmi/mo#isInstanceOf> ;
    owl:someValuesFrom <http://uth.tmc.edu/sbmi/mo#Modality>
  ] ,
  [ rdf:type owl:Restriction ;
    owl:onProperty <http://uth.tmc.edu/sbmi/mo#hasSynchronicity> ;
    owl:qualifiedCardinality "1"^^xsd:nonNegativeInteger ;
    owl:onClass <http://uth.tmc.edu/sbmi/mo#Synchronicity>
  ] ,
  [ rdf:type owl:Restriction ;
    owl:onProperty <http://uth.tmc.edu/sbmi/vmo#hasDistributionMethod> ;
```



```

    owl:qualifiedCardinality "1"^^xsd:nonNegativeInteger ;
    owl:onClass <http://uth.tmc.edu/sbmi/mo#Distribution_Method>
  ] ;
  dc:description "\"The communication channel comes in a variety of forms
    such as face-to-face communication, phone, e-mail, and public speech
    . Choosing communication channels is an integral part of the entire
    course of misinformation transmission.\" Zhou and Zhang, 2007" .

# http://uth.tmc.edu/sbmi/mo#Synchronicity
<http://uth.tmc.edu/sbmi/mo#Synchronicity> rdf:type owl:Class ;
  dc:description "\"representing whether the sender and receiver are
    exchanging information simultaneously\" (Zhou and Zhang, 2007)" .

# http://uth.tmc.edu/sbmi/mo#Modality
<http://uth.tmc.edu/sbmi/mo#Modality> rdf:type owl:Class ;
  dc:description "\"representing how misinformation is encoded (e.g., in
    text, audio, video, image, or any of their combinations);\" (Zhou
    and Zhang, 2007)" ;
  rdfs:comment "form in which misinformation is expressed" .

# http://uth.tmc.edu/sbmi/mo#Distribution_Method
<http://uth.tmc.edu/sbmi/mo#Distribution_Method> rdf:type owl:Class ;
  dc:description "\"representing whether the sender and the receiver are
    physically located at the same place or are distributed while
    sending/receiving misinformation\" (Zhou and Zhang, 2007)" .

# http://uth.tmc.edu/sbmi/mo#Recordability
<http://uth.tmc.edu/sbmi/mo#Recordability> rdf:type owl:Class ;
  dc:description "\"representing whether or not misinformation is archived
    for later reference.\" (Zhou and Zhang, 2007)" .

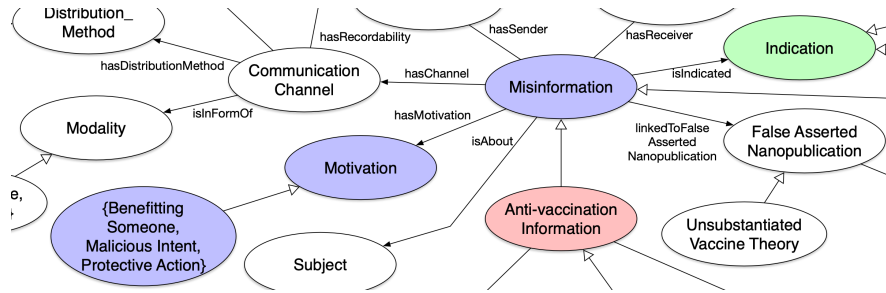
```

**Listing 3.5:** Turtle syntax of Communication Channel, Synchronicity, Modality, Distribution Method, and Recordability concepts from VAXMO

### 3.2.6 CUES TO MOTIVATION AND INDICATION CONCEPTS

VAXMO also represents motivations and cues of misinformation. The Indication concept is used to classify what type of evidence supports misinformation and the Motivation concept that classifies

the reason behind the misinformation content.



**Figure 3.14:** Class structure of VAXMO showing the Motivation concept and also the Indication class. Listing 3.6 shows of the Turtle syntax representation.

Indication concept is subclassed by various anti-vaccine subclasses. The Anti-vaccine evidence has various subclasses (36 classes) that are based on research by Kata's content analysis of various sources.

```
# http://uth.tmc.edu/sbmi/mo#Motivation
<http://uth.tmc.edu/sbmi/mo#Motivation> rdf:type owl:Class ;
dc:description "The production and dissemination of misinformation are
commonly driven by motivations of creating conditions for positive
attributions of worth (Rue, 1994)" (Zhou and Zhang, 2007)" .

# http://uth.tmc.edu/sbmi/mo#Indication
<http://uth.tmc.edu/sbmi/mo#Indication> rdf:type owl:Class ;
dc:description "\"indicators of misinformation\" (Zhou and Zhang, 2007)"
;
rdfs:comment "Zhou and Zhang have stated that this concept would need
further development in the future. \"The definition of a class can
be enriched with additional ontological information. We would like
to leave such extensions for future research.\" (Zhou and Zhang,
2007)" ,
"aka \"cues\" of misinformation" ,
"replace with Provenance ontology" ,
"supporting information for the Content (misinformation)" .

# http://uth.tmc.edu/sbmi/mo#NonImmediacy
<http://uth.tmc.edu/sbmi/mo#NonImmediacy> rdf:type owl:Class ;
rdfs:subClassOf <http://uth.tmc.edu/sbmi/mo#Indication> .
```

```

# http://uth.tmc.edu/sbmi/mo#NonImmediacy
<http://uth.tmc.edu/sbmi/mo#NonImmediacy> rdf:type owl:Class ;
  rdfs:subClassOf <http://uth.tmc.edu/sbmi/mo#Indication> .

# http://uth.tmc.edu/sbmi/mo#Response_Latency
<http://uth.tmc.edu/sbmi/mo#Response_Latency> rdf:type owl:Class ;
  rdfs:subClassOf <http://uth.tmc.edu/sbmi/mo#Indication> .

# http://uth.tmc.edu/sbmi/mo#Self-Reference
<http://uth.tmc.edu/sbmi/mo#Self-Reference> rdf:type owl:Class ;
  rdfs:subClassOf <http://uth.tmc.edu/sbmi/mo#Indication> .

# http://uth.tmc.edu/sbmi/ontology/vaxmo#Anti-vaccine_Evidence
<http://uth.tmc.edu/sbmi/ontology/vaxmo#Anti-vaccine_Evidence> rdf:type
  owl:Class ;
  rdfs:subClassOf <http://uth.tmc.edu/sbmi/mo#Indication> ;
  rdfs:comment "derived from Kata, 2010." .

# http://uth.tmc.edu/sbmi/ontology/vaxmo#No_Evidence
<http://uth.tmc.edu/sbmi/ontology/vaxmo#No_Evidence> rdf:type owl:Class ;
  rdfs:subClassOf <http://uth.tmc.edu/sbmi/mo#Indication> ,
    [ rdf:type owl:Restriction ;
      owl:onProperty <http://uth.tmc.edu/sbmi/ontology/vaxmo#hasNumberOfCues
        > ;
      owl:qualifiedCardinality "0"^^xsd:nonNegativeInteger ;
      owl:onDataRange xsd:int
    ] .

```

**Listing 3.6:** Turtle syntax of Motivation and Indication from VAXMO

### 3.2.7 VACCINE THEORIES CONCEPTS

Before further discussing this aspect of VAXMO, we need to introduce an RDF model for assertions, called Nanopublications. Nanopublications are micropublishing format to annotate scientific assertions, like mosquitos spread malaria to borrow an example from a certain study ([Kuhn et al., 2015](#)). Nanopublications are essentially an RDF graph model for one proposition or triple that state

a scientific finding. This graph model also annotates authoring information, like the name of the researcher who discovered this finding, when it was discovered, etc. (Groth et al., 2013) There is also a unique identifier associated with assertion for querying or machine reasoning (Groth et al., 2010). Code Listing 3.7 shows an adapted example by Groth and associates (Groth et al., 2013).

```
@prefix : <http://example.org/pub1#> .
@prefix ex: <http://example.org/> .
@prefix np: <http://www.nanopub.org/nschema#> .
@prefix prov: <http://www.w3.org/ns/prov#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

# HEADER SECTION
:head {
  ex:pub1 a np:Nanopublication .
  ex:pub1 np:hasAssertion :assertion .
  ex:pub1 np:hasProvenance :provenance .
  ex:pub1 np:hasPublicationInfo :pubInfo .
}

# ASSERTION SECTION
:assertion {
  ex:trastuzumab ex:is-indicated-for ex:breast-cancer .
}

# PROVENANCE SECTION
:provenance {
  :assertion prov:generatedAtTime "2012-02-03T14:38:00Z"^^xsd:dateTime .
  :assertion prov:wasDerivedFrom :experiment .
  :assertion prov:wasAttributedTo :experimentScientist .
}

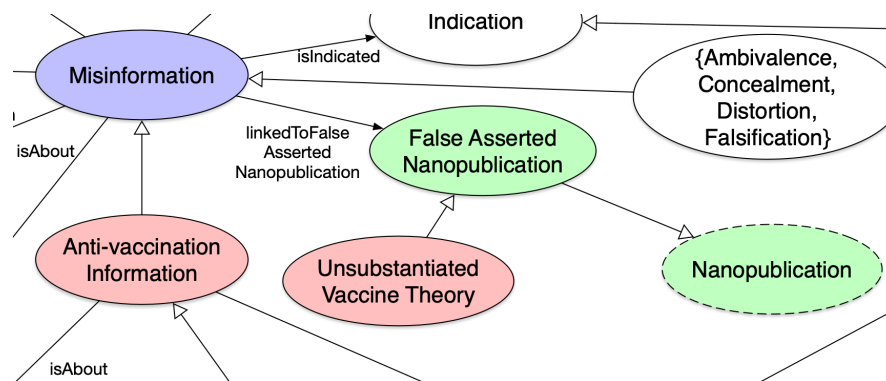
# PUBLICATION SECTION
:pubInfo {
  ex:pub1 prov:wasAttributedTo ex:paul .
  ex:pub1 prov:generatedAtTime "2012-10-26T12:45:00Z"^^xsd:dateTime .
}
```

**Listing 3.7:** Sample nanopublication encoding adapted from Groth et al. (2013)

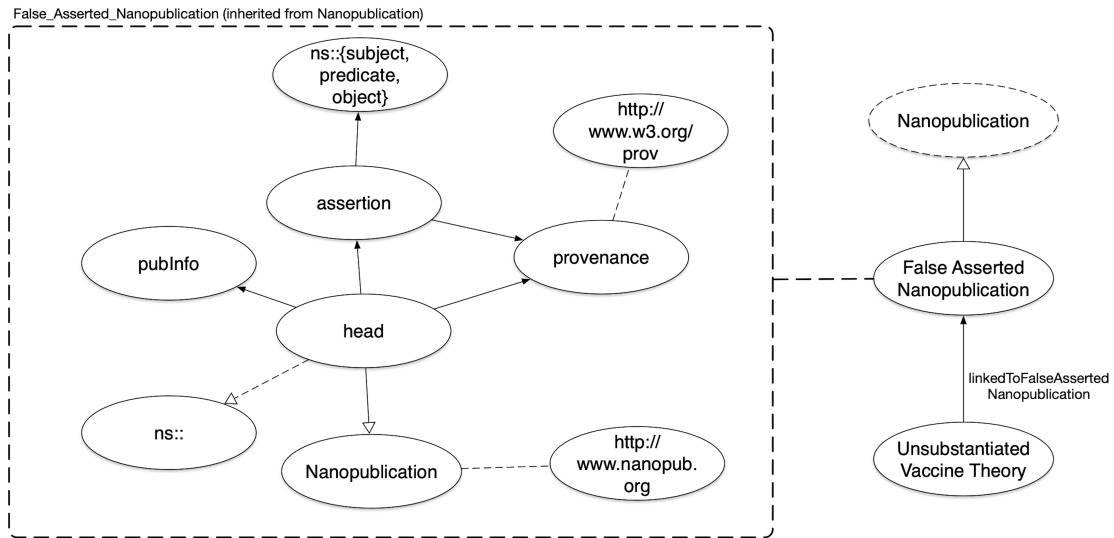
The nanopublication above is coded in RDF with four basic components that outline its structure

- header, assertion, provenance, and authoring information. The header part declares annotation labels of the nanopublication. The assertion of the nanopublication encodes the scientific assertion. In the example above, the assertion is “vaccines were developed for population control”. The assertion stated is incorrect, but as we will explain later, the nanopublication’s model to annotate false assertions serves a purpose for VAXMO. The provenance and authoring information annotates some origin information for reference purposes of the assertion.

In Figure 3.15 and 3.16, we created a “Vaccine Theories” concept that inherits the nanopublication structure and linked to Content to provide an association with Anti-Vaccination information. Vaccine Theories is the concept in VAXMO that is used to annotate a proposition or triple like “vaccine causes seizures”. In the example, *vaccine* → *causes* → *seizures* (Figure 3.16) is the false assertion annotated as Anti-vaccination information. This false assertion abstracted by the nanopublication model, provides information on the origin of the assertion by where and who it originates from using the Provenance ontology. Overall, the figure is a graphical representation of the code listing that shows the link to a concept that associates it with anti-vaccination.



**Figure 3.15:** Class structure of VAXMO showing False Asserted Nanopublication concept that is a subclass of Publication. Unsubstantiated Vaccine Theory is a subclass of False Asserted Nanopublication that annotates triples of vaccine misinformation.



**Figure 3.16:** Further details of the nanopublication structure of the False Asserted Nanopublication class.

### 3.2.8 VAXMO'S METRICS

The total number of classes from The Vaccine Misinformation Ontology amounted to 116, with 6 data properties and 20 object properties. The latest version of VAXMO is hosted at the SBMI Ontology Research Group website and was authored using Protégé using OWL2 syntax. In Chapter 5, we review the quality evaluation of VAXMO. In Chapter 6, we introduce some use cases that are outside of the utilization by the conversational agent for HPV vaccine in Chapter 6.

### 3.3 SUMMARY

The contents of this chapter examined our work in developing ontological knowledge bases for health consumer-centric vaccine knowledge. The knowledge structures between consumers and experts differ in how they represent their notions regarding the vaccine, and this would be assumed to be true for the HPV vaccine. This understanding led us to develop the Vaccine Information Statement Ontology (VISO), and its extension Vaccine Information Statement Ontology for HPV Vaccine (VISO-HPV). Both of these ontologies derive their information from Vaccine Information Statements, which are federally mandated informed consent documents to given to patients before their vaccination (as required by National Childhood Vaccine Injury Act). Vaccine misinformation is an important knowledge detail to cover with health consumers, and this led to the development of the Vaccine Misinformation Ontology (VAXMO) that is built on the Misinformation Ontology by Zhou and Zhang. Overall, these health consumer ontologies will support the domain knowledge layer of our proposed conversational agent for the HPV vaccine. VISO, VISO-HPV and VAXMO are all available at the SBMI Ontology Research Group website <sup>†</sup>.

Chapter 2 also reviewed our work in the area of ontology evaluation which is an important facet of knowledge engineering to validate the aforementioned ontologies. Some of the contributions in this area include creating a benchmark standard based on a library of ontologies and an automated web-based tool called OntoKeeper. This work will play important role in validating the quality of the ontologies, which is covered in Chapter 5.

We would like to thank Dennis Wang for his work and contribution on VISO-HPV.

---

<sup>†</sup><https://sbmi.uth.edu/ontology/>

*Pay no attention to that man behind the curtain! The  
Great OZ has spoken!*

The Wizard

# 4

## Interaction-driven application ontologies

AN ASPECT of this work involves understanding the user and designing the necessary tools based on the users' response. For our endeavor, we need to understand how the user would respond to a conversational agent for HPV vaccine (since no such tool exists) and from that knowledge, engineer the appropriate methods and technology. The lack of similar conversational agents led us to implement a simulated version of the agent. This would permit us not only to understand the needs of the users, but also gain an insight how to best develop the conversational agent by collecting data to support the advancement of our work, which has led to developing the technology for interaction. In this chapter, we introduce a simulation experiment involving a conversational agent for HPV vaccine, and reveal how some of the data collected influenced the design of the

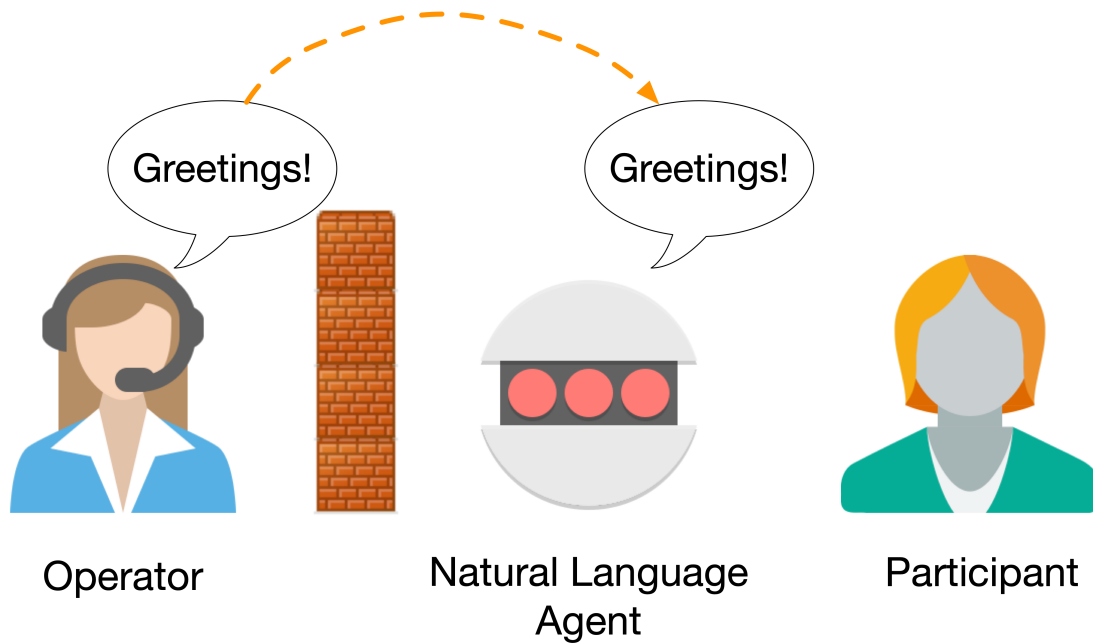


interaction technology for dialogue (Section 4.2). Also, we will describe the collaborative work with a colleague in developing a unique visual interface that can be integrated with agents.

#### 4.1 WIZARD OF OZ EXPERIMENT

The Wizard of OZ experiment is said to take its inspiration from the famous novel and its movie adaption of its same name. Other names for this experiment may have been “Pay No Attention to the Man Behind the Curtain” (Fraser & Gilbert, 1991). The Wizard of OZ (WOZ) experiment simulates the speech interaction between a prospective user of a natural language interface system - manifested as software application agent or robot. The modality of the interface may either be text or speech emanating from the agent. The interaction of the natural language interface is driven by a “wizard” or “operator” providing the dialogue of the agent (Figure 4.1). This would give the user the appearance of an automated agent, thereby, yielding a genuine response of the user towards the agent. The general purpose of this experiment is to forecast, understand, and collect data of the interaction with users in order to build a completely automated version of the agent.

There are several benefits for conducting WOZ experiments for automated natural language interface agents (Cohen et al., 2004). One is the potential to do early testing, either for specific software components that will be integrated into the automated agent or the dialogue uttered by the agent. Speech recognition systems may lack the sufficient language models to interpret the utterance of the users which may cause recognition failures and impede the perception of the user towards the agent. Since the simulated system is coordinated by a human user behind the scenes, there is the benefit of vocabulary coverage. This may also be a disadvantage as some users may not believe that automated agent would have a high vocabulary coverage and may dissuade the belief in the agent’s authenticity. Another benefit is the relative ease in running and updating to the WOZ versus making updates to an automated prototype.



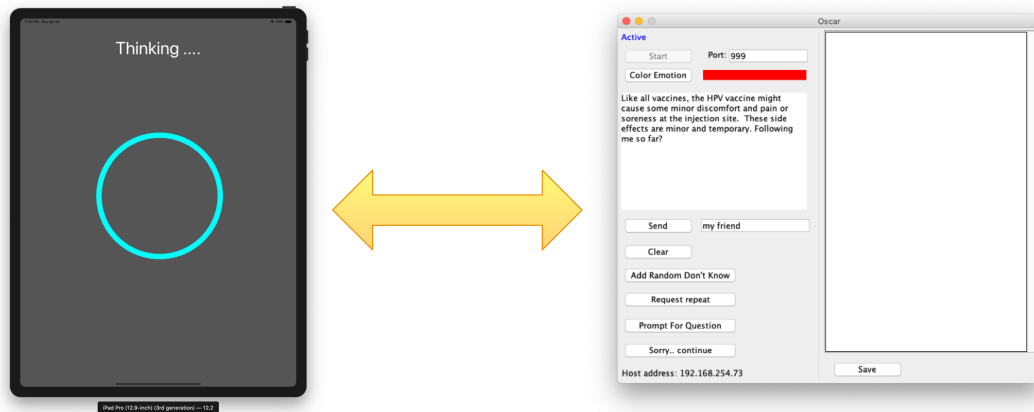
**Figure 4.1:** Wizard of OZ experimental protocol. "Support Icon" (CC Attribution 3.0) by Squid Ink (Squid Ink, 2019), "Brick wall Icon" (CC Attribution 4.0) by Anna Shlyapnikova (Shlyapnikova, 2019), "Gnome robots icon" (GNU General Public License v3.0) by Papirus Development Team (Papirus Development Team, 2019), and "User female alt Icon" (Public Domain license) by paomedia (paomedia, 2019)

Generally for the WOZ experiment to be effective, Fraser and Gilbert outline a few prerequisites (Fraser & Gilbert, 1991).

1. *The simulation must be possible* - The experiment has to be feasible to mimic the automated version.
2. *The future system must be specifiable* - Researchers need to know how the system will interface with the participant.
3. *The simulation must be convincing* - The experiment is realistic or believable as a machine to the user.

#### 4.1.1 TOOLS AND METHODS

To understand the user experience of a conversational agent for HPV vaccine and attain guidance on how to develop the conversational agent, we embarked on conducting a Wizard of OZ experiment on users who may be decision makers for the HPV vaccine - either for themselves or their children. At the time of my research, no dialogue system tool for HPV vaccine existed. The primary foci for executing the Wizard of OZ technique was to discover how to converse with and counsel users on the HPV vaccine, gain insight on the technology to be implemented in the interaction aspect of this tool, and collect data. Later in Chapter 5, we will discuss some of the user experience data from HPV vaccine decision makers – parents and young college adults.



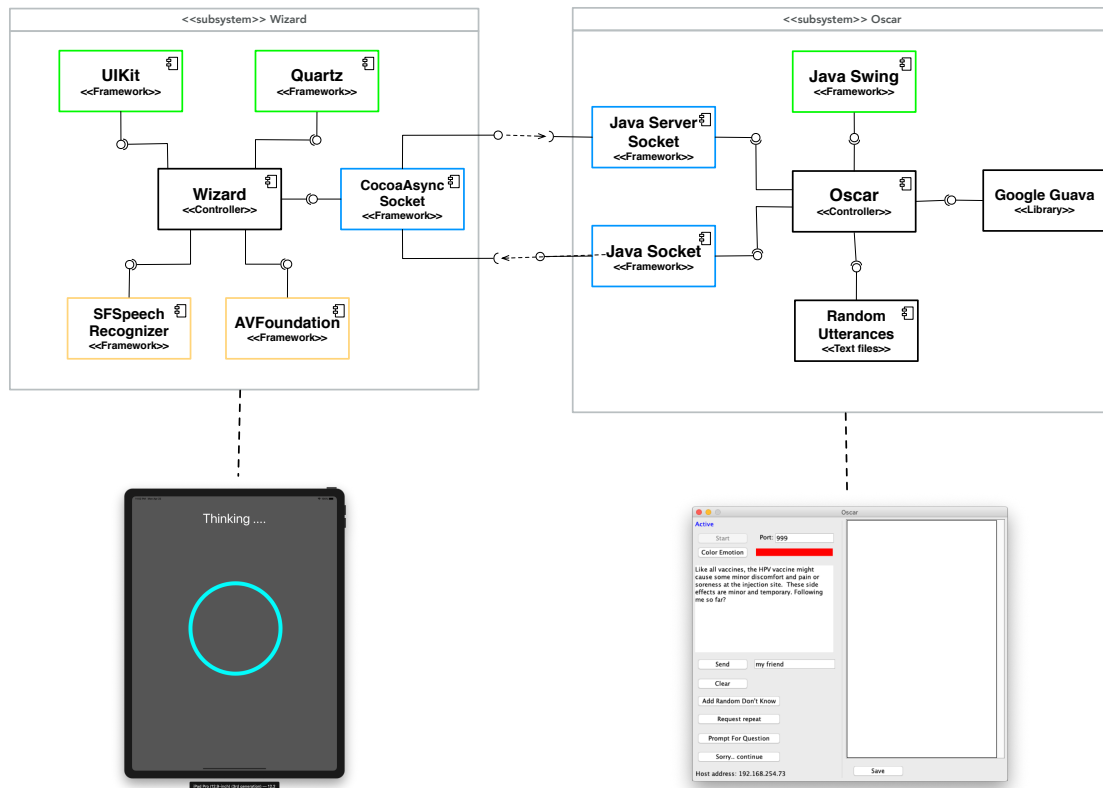
**Figure 4.2:** Simulation software for the study’s Wizard of OZ experiment. Left application is the natural language speech interface (“Wizard”) and at the right is the remote controller (“Oscar”) for “Wizard”

**SIMULATION SOFTWARE** The deployment for the conversational agent was envisioned to be a tablet-based kiosk-like device in the waiting area of a clinic. I developed a basic iOS application for an Apple iPad Pro that allowed for hands-free speech interaction with the device (“Wizard”). The application was developed using Apple’s XCode and iOS SDK (v12) utilizing various software

libraries like *SFSpeechRecognition* (speech recognition) and *AVFoundation* (text to speech). For convenience, the voice profile used was the default voice for Apple's Siri application. Complementing the tablet was a Java-based desktop application ("Oscar") that can be used on any machine with a JVM (Java 8). This application was designed for the operator to communicate to the user to masquerade as the tablet. The software notifies the operator when the user is ready to use the tablet, and this will signal the operator to begin the conversation. A text box serves as the space for the operator to type or copy and paste the messages to be uttered by the user. Optionally, several buttons were available to send randomly canned responses like variations to prompt the user if they have a question (e.g. "I am open to any questions you might have [name]. Have any questions?", "Would you like to ask a question [name]?"). Operators can create a set of variation of utterances that express the same notion and the software will randomly pick a variant. This saves time for the operator to type or think about a response. The software can collect dialogue exchanges between the participant and the simulated agent and display it on the panel, and the operator can save the chat log.

**DIALOGUE SCRIPT** One of the goals of this research is to not only inform the user about the HPV vaccine, but also to have some impact on their health beliefs towards the HPV vaccine. As a standard baseline, we utilized the Carolina HPV Immunization Attitudes and Belief Scale (CHAIS), which is a validated survey to measure the health beliefs of the HPV vaccine among parents (McRee et al., 2010a; Gowda et al., 2012). Variations of CHAIS has been developed for young adults and has been shown to have the same validation (Dempsey et al., 2014).

The CHAIS survey served as a touchstone for talking points about the HPV vaccine. Each notion for the survey questions was expressed as dialogue that communicated that piece of information. For example, the survey question, "How effective do you think the HPV vaccine is in preventing cervical cancer?". This particular question expressed the effectiveness of the HPV vaccine to prevent

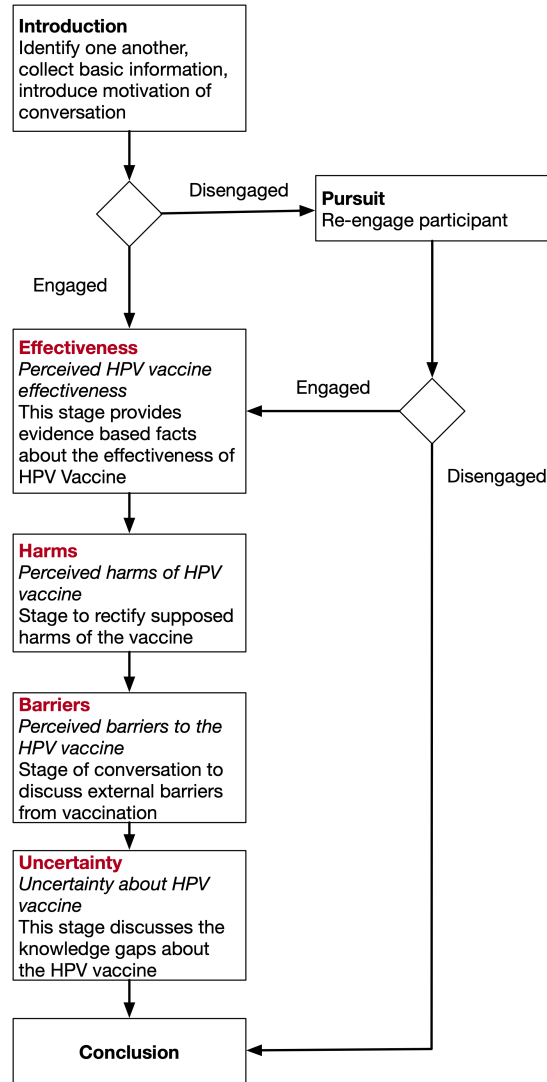


**Figure 4.3:** A basic UML component diagram abstracting this project’s Wizard of OZ software.

cancer of the cervix. The dialogue for this notion became: “If your child is vaccinated with the HPV vaccine it will protect against various HPV viruses which causes many precancerous and cancerous lesions in males and females” and followed by an inquiry if they understood the information (e.g. “Following me so far, [name]?”). The survey was also segmented into categories related to the health belief model - Harms, Barriers, Effectiveness and Certainty. Harms had six questions/talking points, Barriers had five, and Effectiveness and Certainty had two and three respectively.

After the completion of the script in the form of simple state diagram, it was reviewed by public health researchers \* and a pediatrician (Julie Boom, MD). Complex information or information that

\*Rebecca Cunningham, MPH of Texas Children’s Hospital, Lara S. Savas, PhD of UTHealth School of Public Health, and Laura A. Shay, PhD of UT-San Antonio’s School of Public Health



**Figure 4.4:** General framework for discussing vaccine.

would invite a personal detailed discussion was minimized. For example, number of doses could invite the user to ask which series of the HPV vaccine they should take - two-dose series or three-dose series. Any discussion of the series would involve knowing when the health consumer had the vaccine, what their age, etc. The dialogue flow also permitted opportunities for the user to ask a question at any time. To closely represent how the agent would respond to participant questions,

we used the knowledge triples from VISO-HPV (discussed in Chapter 3). If the question was out of scope of the knowledge base, the operator would respond with an apology and a recommendation to ask their provider (e.g. “[name] That’s a good question. Sadly, I am not sure how to answer that question. I would recommend writing down that question and asking your doctor when you next see him or her. Shall I continue?”), along with a follow up if they have another question. Another feature of the dialogue was the inclusion of a pursuit flow (contributed by Rebecca Lin). Provider pursuit is a recommendation tactic detailed by Opel et al. (2013) where if a patient is disinterested in the HPV vaccine to verbally nudge the patient towards having an interest in the vaccine. Research have shown that pursuing vaccine hesitant parents encourage the parent to accept the vaccination recommendation (Opel et al., 2012, 2013). Figure 4.4 shows the overall architecture of the dialogue. We performed WOZ experiment on two occasions, a preliminary trial with parents with at least one child under 18<sup>†</sup> and another test with young adults at Texas A & M University at College Station<sup>‡</sup>. Detailed discussion of the experiments is found in Chapter 5. The chat logs from the trials and the dialogue script influenced the design of an ontology in Section 4.2.1 Patient Health Information Dialogue Ontology and the chat logs would later be used to test the automated dialogue system (Chapter 5).

## 4.2 ONTOLOGY-DRIVEN DIALOGUE ENGINE

Using ontologies for dialogue management specifically in the health-related domain offers some unique opportunities in healthcare research for conversational agents. According to Kennedy et al. (2012) and demonstrated in research (Bickmore et al., 2011, 2013), there is potential to base

---

<sup>†</sup>The University of Texas Health Science Center’s Committee for the Protection of Human Subjects approved this study (HSC-SBMI-17-0533)

<sup>‡</sup>The University of Texas Health Science Center’s Committee for the Protection of Human Subjects approved this study (HSC-SBMI-19-0102), Texas A & M University Human Research Protection Program (IRB2019-0118M)

electronic-based interventions in established health behavioral models like the trans-theoretical model (Prochaska & DiClemente, 2005), motivational interviewing (Miller & Rollnick, 1991), or the health belief model (Janz & Becker, 1984; Becker, 1974). As we had mentioned in Chapter 2, there is very little research in utilizing ontologies for dialogue systems for health yet there is unrealized potential to exploit them from this area.

Coming off the heels of some of the initial work completed in the Wizard of OZ experiment, we developed an application ontology that can provide the facilities to manage dialogue flow and the contextual dialogue information (e.g. whether a certain utterance has been evoked). This application ontology could also be used in other similar domains that involve health information communication, and also be shared and ported to various other systems – like embodied conversational agents or web-based agents. As we have noted earlier, this ontology could be integrated with models of health behavior. If an ontology is harnessed for health communication, there is the potential to represent and structure the complexity of communicating health information and knowledge, and if that is accomplished there is the potential to share that ontology to other systems and establish a formalization and consistent standards in the domain of communicating health information to patients.

This section of the chapter will outline the design of PHIDO (Patient Health Information Dialogue Ontology) and the development of the software controller that is directed by PHIDO. The PHIDO class hierarchy is represented by four basic classes - *Utterance* class, *Speech Task* class, *Goal* class and *Discussion* class - to be discussed in the subsequent sections. A later section will introduce the dialogue flow transition that relies on the PHIDO model.



## 4.2.1 PATIENT HEALTH INFORMATION DIALOGUE ONTOLOGY

### 4.2.1.1 UTTERANCE CLASS

The Utterance class is the basic building block for PHIDO as it defines information related to evoking an utterance between the agent and the health consumer user. The basic Utterance class supports data types that enable the machine to facilitate utterances that are exchanged between it and the human users. These data types include:

- *hasUtterancePriority*, an integer type, to indicate the utterance's rank. This property enables the agent to organize and select utterances to be evoked.
- *hasUtteranceString*, a string type, that is the actual verbiage to be spoken or printed by the agent.
- *hasBeenSaid*, a boolean type, to indicate whether a specific instance of the utterance has been spoken or printed by the agent.
- *hasUtteranceExamples*, a string type, similar to the *hasUtteranceString*, but used by the agent to discern utterances that are not evoked by the agent (i.e. Participant Utterance - see below)
- *hasFocus*, a boolean type, to designate that the utterance is the current utterance. This property is used for dialogue flow which we will discuss in the subsequent section.

PHIDO provides links with the Utterance class and between the Utterance class and the Speech Task class (discussed in a later section). The object property of *utteranceLink* serves as a link to connect instances of the Utterance class. This object property has several subtypes - *follow* and *precedes*. These two subtypes of object properties are the inverse of each other. If an utterance instance *a* follows another utterance instance *b*, the machine reasoner (e.g. Pellet (Sirin et al., 2007), HermiT (Glimm et al., 2014), etc.) would reveal *b* precedes *a*.

PHIDO has several subclasses of the Utterance class. The essential subclasses of the Utterance class are the System Utterance and Participant Utterance. The System Utterance describes utterance

types that are elicited by the agent. This utterance, based on a review of chat logs and dialogue script, has various subclasses of its own, as shown in Table 4.1. Similarly, Table 4.2, has a set of subclasses of Participant Utterances, which are utterances that are expressed by the user to the agent. These subclasses of the System and Participant Utterance are utilized in some of the Speech Tasks (Section Speech Task).

In a minimal effort to support upper level ontologies, like the Basic Formal Ontology (BFO) (Smith & Grenon, 2019), that unites the hierarchical organization of biomedical ontologies, we incorporated Searle's speech classification (Searle, 1976). Currently, BFO has one class concept called "utterance", which could be subclassed by high-level classifications for speech. Searle's speech classification are high-level categories that describe utterances. These classifications include:

- "Assertive: committing the speaker to something's being the case (suggesting, putting forward, swearing, boasting, concluding)
- Directives: attempts by the speaker to get the addressee to do something (asking, ordering, requesting, inviting, advising, begging)
- Commissives: committing the speaker to some future course of action (promising, planning, vowing, betting, opposing)
- Expressives: expressing the psychological state of the speaker about a state of affairs (thanking, apologizing, welcoming, deploring)
- Declarations: bringing about a different state of the world by the utterance (including many of the performative examples above; we resign, you're fired)" (Searle, 1976)

While BFO serves as a template for biomedical and health ontologies with canonical information, the PHIDO ontology is geared to be an application ontology for supporting software tasks. Nonetheless, work towards integrating utterance types would need to be investigated in later endeavors.

**Table 4.1:** System Utterance classes

System Utterance	Speech act type
Acknowledgment	Expressive
Agenda	Commissive
Answer	
<i>No Answer</i>	Representative
<i>Has Answer</i>	Representative
Apology	Expressive
Capitulate	Expressive
Compassionate Utterance	
<i>Condolence</i>	Expressive
<i>Happy For</i>	Expressive
Confirm Health Information	Directive
System Declaration	
<i>Disclaimer</i>	Representative
<i>Topic Transition</i>	Commissive
System Farewell	
<i>Concluding Farewell</i>	Expressive
System Greet	Expressive
Inform	Representative
Inquire Personal	Directive
Interview Question	Directive
Option	
<i>Clarification Options</i>	Directive
<i>Question Options</i>	Directive
<i>Topic Options</i>	Directive
Overview	Commissive
Request	Directive
Request Repeat	Directive
Satisfaction Prompt	Directive
System Declaration	
<i>Disclaimer</i>	Representative
<i>Topic Transition</i>	Commissive
System Introduction	Representative

**Table 4.2:** Participant Utterance classes

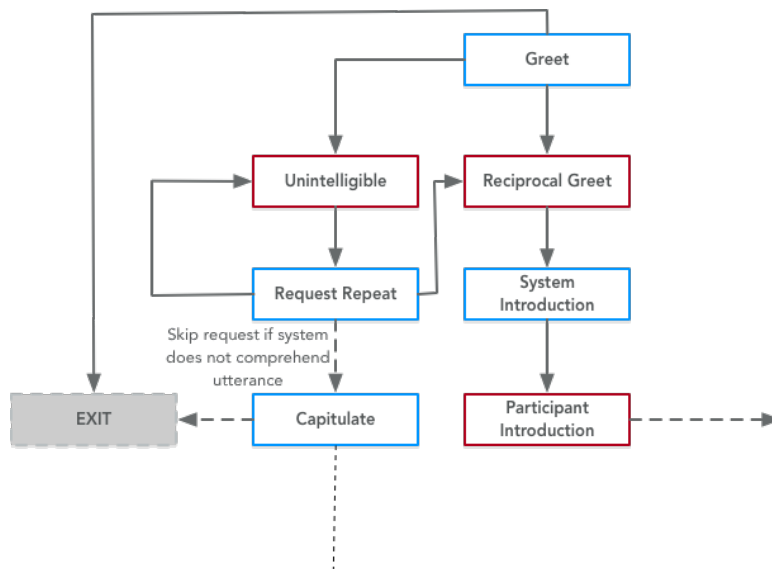
Participant Utterance	Speech act type
Acceptance	Expressive
Negative Utterance	
<i>Disconfirmation</i>	Expressive
<i>Negative Personal Status</i>	Representative
Participant Farewell	Expressive
Participant Introduction	Assertive
Personal Status	
<i>Negative Personal Status</i>	Expressive
<i>Positive Personal Status</i>	Expressive
Positive Utterance	
<i>Confirmation</i>	Representative
<i>Positive Personal Status</i>	Expressive
Prattle	<i>NA</i>
Question	
<i>Divergent Question</i>	Directive
Reciprocal Farewell	Expressive
Reciprocal Greet	Expressive
Request System	
<i>Request System Repeat</i>	Directive
Unintelligible	<i>NA</i>

#### 4.2.1.2 SPEECH TASK CLASS

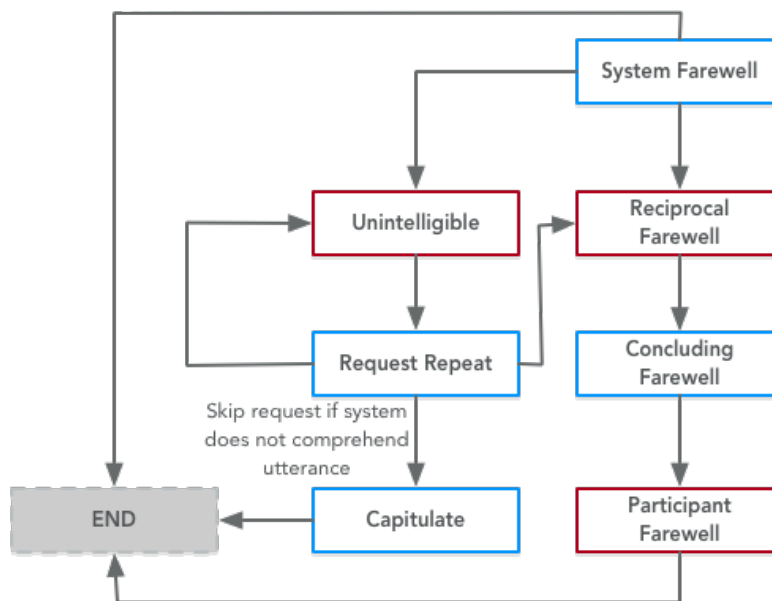
PHIDO contains a class concept Task which has a subclass Speech Task. Essentially a group of Utterances associated with a Speech Task defines the meaning and purpose of the Speech Task. Many instances of an Utterance class is linked to a Speech Task with *belongsToASpeechTask*, and every Speech Task has many instances of the Utterance class (*hasUtterance*). Similar to the Utterance class, the Speech Task yields various subclasses - Pleasantry Task (Salutation and Valediction), Proposition Task (Initiate Discussion, Transition to Topic, Interview Participant, Discuss Health Topic), and Question and Answering Task. Figures 4.5 through 4.11 display all of the Speech Tasks that are supported in PHIDO, with the class definitions encoded. In all of the Speech Task classes there are various System and Participant Utterances “chained” together to help the agent direct the flow of the dialogue. For brevity sake, Table 4.3 summarizes each of the Speech Tasks.

**Table 4.3:** Outline list of the various Speech Tasks and their function/purpose.

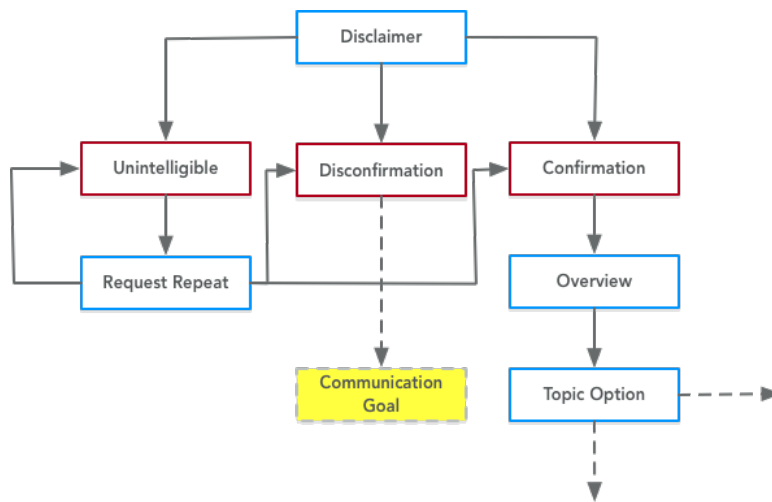
Class	Parent Class	Function
Salutation Task	Pleasantry Task	Formal exchange of greetings with the user at the start of discussion
Valediction Task	Pleasantry Task	Ends the discussion session with the participant with exchanges of good-byes
Initiate Discussion	Proposition Task	Formal introduction of the theme of the counseling
Transition to Topic	Proposition Task	Sequence to change topic of the conversation
Interview Participant	Proposition Task	Probe for information about the user; engage in “small talk” with user
Discuss Health Topic	Proposition Task	Communicate and confirm health information to user
Question and Answering	Speech Task	Sequence to engage and answer the a question from the user



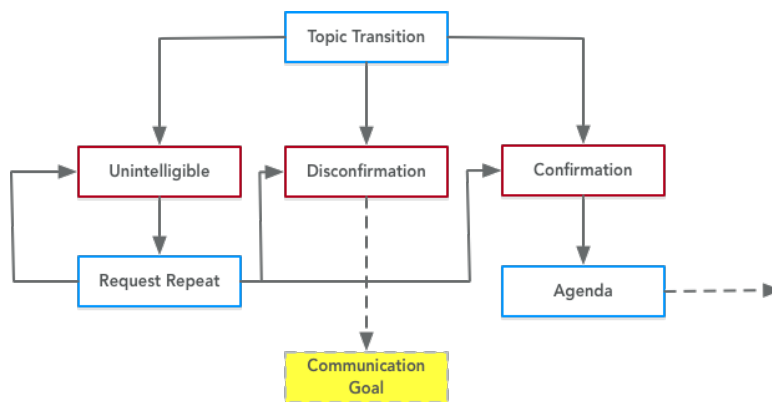
**Figure 4.5:** The Salutation class from PHIDO. Blue is the System Utterance and red is the Participant Utterance



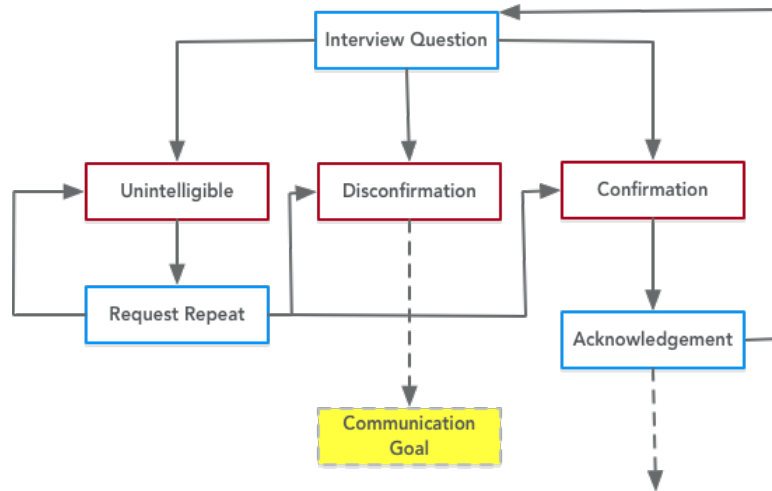
**Figure 4.6:** The Valediction class from PHIDO. Blue is the System Utterance and red is the Participant Utterance.



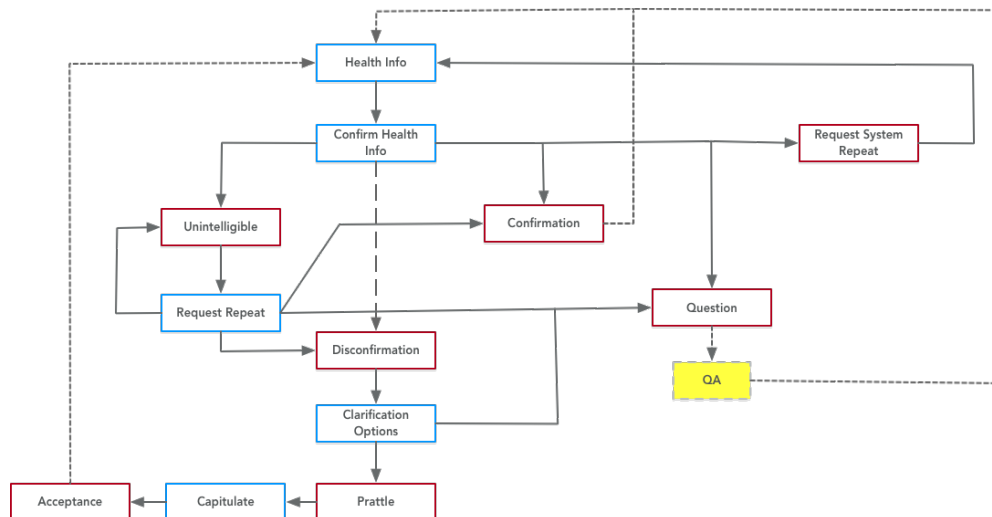
**Figure 4.7:** The Initiate Discussion class from PHIDO. Blue is the System Utterance and red is the Participant Utterance.



**Figure 4.8:** This Transition to Topic class from PHIDO. Blue is the System Utterance and red is the Participant Utterance.

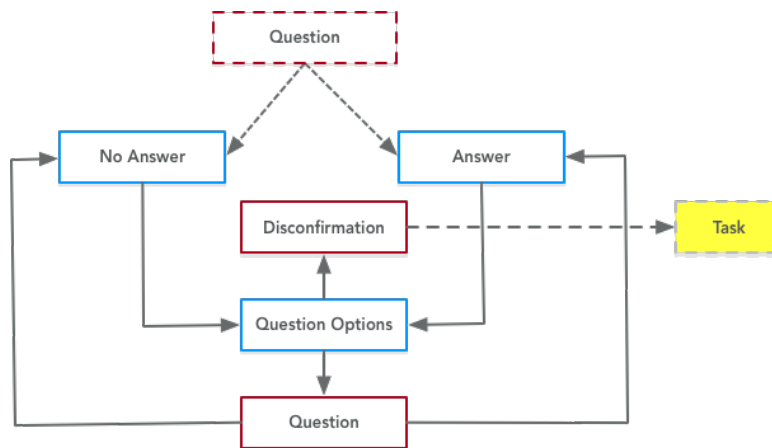


**Figure 4.9:** The Interview Participant class from PHIDO. Blue is the System Utterance and red is the Participant Utterance.



**Figure 4.10:** The Discuss Health Topic class from PHIDO. Blue is the System Utterance and red is the Participant Utterance.





**Figure 4.11:** The Question Answering Task class from PHIDO. Blue is the System Utterance and red is the Participant Utterance.

#### 4.2.1.3 DISCUSSION AND COMMUNICATION GOALS CLASSES

PHIDO has a couple of high level classes that organize and align the Utterance and Speech Task classes. The Communication Goal class collates the various Speech Task class instances to form an overall objective. These Communication Goals could potentially be aligned with health behavior theories, like the health belief model discussed earlier in this chapter. Figure 4.12 shows the association that is encoded in PHIDO where a Communication Goal has one-to-many Speech Task class(es). Also the Communication Goal has several subclasses relating to the health belief model - Communicate Benefits, Communicate Uncertainty, Communicate Effectiveness, Communicate Harms. Overall, this alignment assist in the organization and traceability of the communication with the health belief model constructs. Lastly, every communication goal belongs to a general Discussion class. Potentially, PHIDO can encode and serve as repository to various types of communication efforts - health communication for diabetes management, genetic counseling, and cancer survivorship counseling, etc.

Patient Health Information Dialogue Ontology was published in OWL2 using Protégé (Musen et al., 2015) (Figure 4.13). PHIDO has 86 classes, 9 object properties and 5 data properties. PHIDO is available at the SBMI Ontology Research Group page.

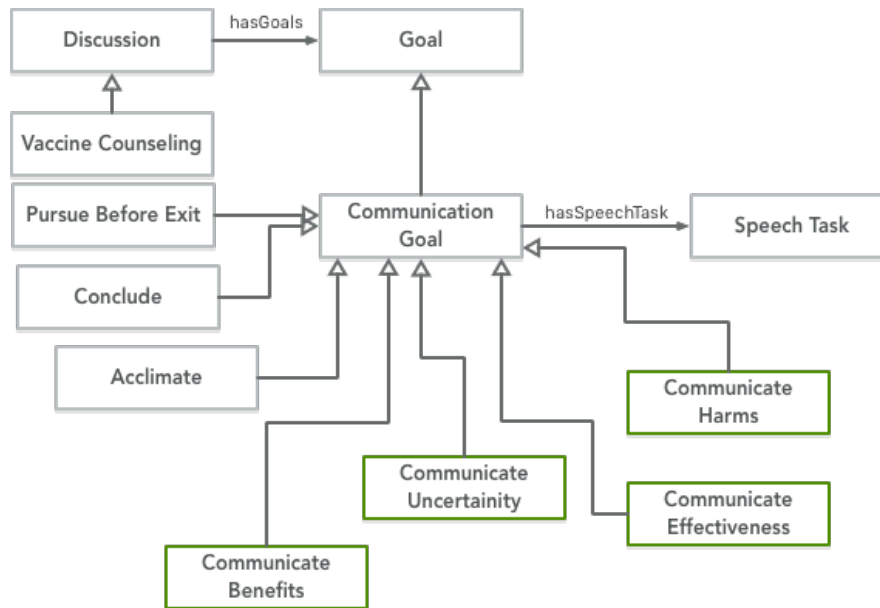


Figure 4.12: UML class diagram of the relationship between the Communication Goal class with Speech Task class and the Discussion class.

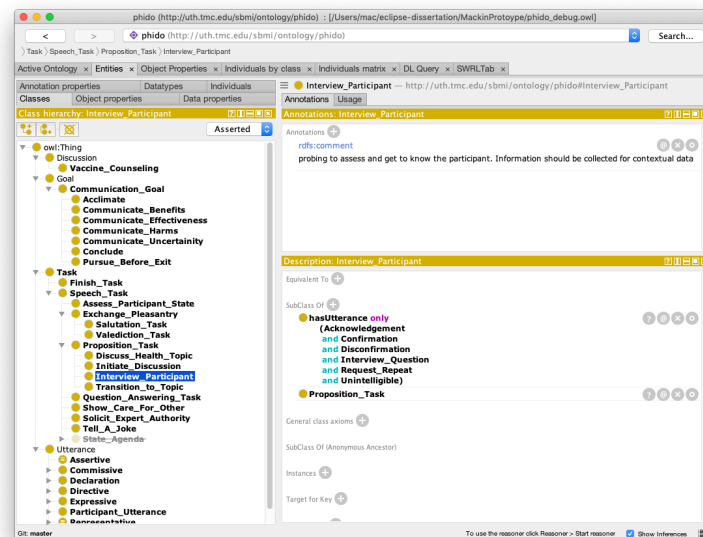
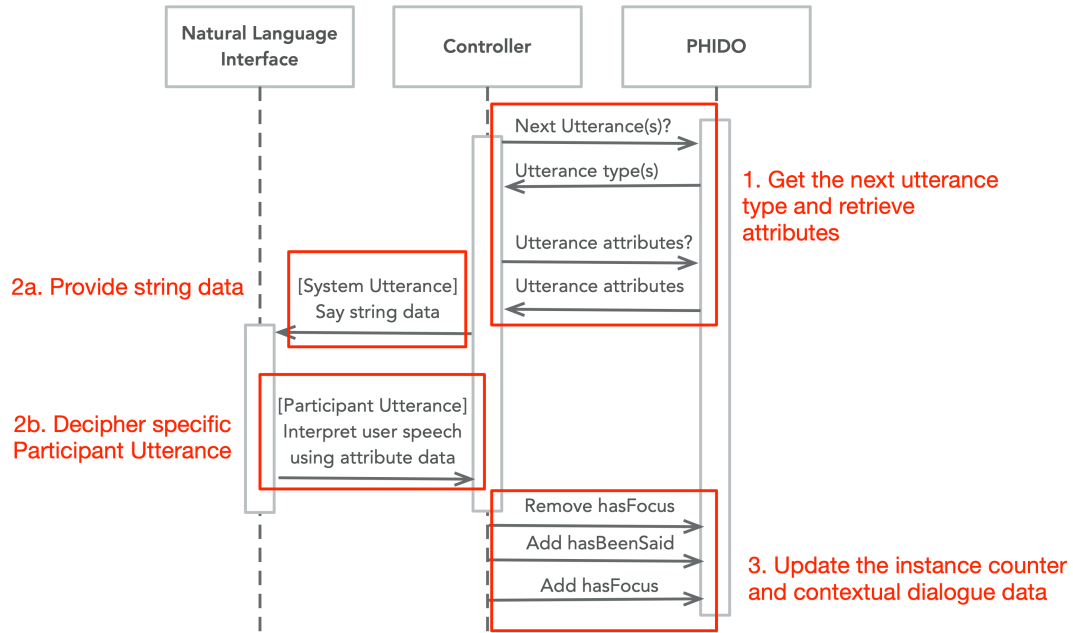


Figure 4.13: Screenshot of Protege with PHIDO.

#### 4.2.2 CONVERSATIONAL ONTOLOGY OPERATOR ENGINE (COO)

The basis of PHIDO is to provide the agent with the structure of dialogue and understand some attributes of utterances and other aspects of conversations. This goal advances the opportunity to develop software components that help agents to interact with users and communicate health information. From this knowledge encoded in the ontology, we can compose an algorithm that utilize PHIDO to guide the system to interact with health consumers.

Figure 4.14 displays the core sequence which involves a natural language interface, the software controller (i.e. software wrapper), and PHIDO. This sequence provides the mechanism for the agent to control and transition the dialogue flow and capture contextual dialogue information to make necessary decisions during its interaction with the user. The sequence starts with an utterance instance with its data property of *hasFocus* set to *true*. From that point the sequence begins by having the software controller query for the next Utterance type and what the next Utterance's attributes. If the next instance utterance is a System Utterance, this would signal the controller to send a text string for the natural language interface to evoke that string data, either by printing it out for the user on the screen or speaking it through text-to-speech component. However, if the next utterance instances is a Participant Utterance, the controller will receive data from the natural language interface to interpret the specific type of Participant Utterance based on attribute information. Afterwards, the contextual dialogue information is updated - setting the current *hasFocus* attribute to *false*, setting the next utterance instance's *hasFocus* attribute to *true*, and optionally setting the previous utterance instance's *hasBeenSaid* to *true*. The sequence then repeats again till the end of the dialogue instances.



**Figure 4.14:** A UML sequence diagram describing the transition mechanism for PHIDO's dialogue flow

#### 4.2.2.1 SOFTWARE ENGINE FOR DIALOGUE INTERACTION

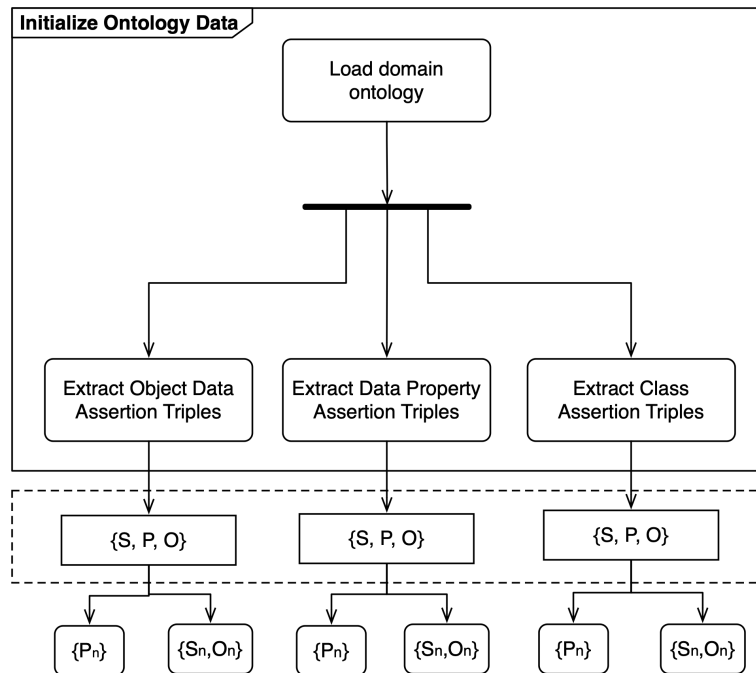
we have developed a software controller that implements the above-mentioned transition sequence called Conversational Ontology Operator (COO). COO was developed using Java 8 and harnesses RDF4j (Eclipse Foundation, 2019) to query and modify the instance level data of the PHIDO ontology. COO also uses the OWL-API (Horridge & Bechhofer, 2011) (5.1.0) and HermiT reasoner library (Glimm et al., 2014) (1.4.1) for supporting functions. In Chapter 5, we describe testing COO using the chat logs from the Wizard of OZ experiments to assess the software engine's ability to handle the dialogue flow.

The Conversational Ontology Operator is also supported by a question answering mechanism called Frankenstein Ontology Question-Answering for User-Centric Systems (FOQUS). This software component was developed in Java 8 and used a combination of OWL-API (5.1.0) (Horridge & Bechhofer, 2011), Stanford Core NLP (3.92) (Manning et al., 2014), MorphaStemmer (Minnen et al.,

2001; Schmitz & Thomson, 2013), HermiT reasoner library (1.4.1) (Glimm et al., 2014) and term and string similarity libraries -Simmetrics-Core (Korstanje, 2019) for various string similarity measures and extjwnl (Autayeu, 2016) for WordNet 3.1 synonymous matches. We also experimented with vector-based similarity using Numberbatch (Speer & Lowry-Duda, 2017), a vector model based on ConceptNet (Speer & Havasi, 2013). Chapter 5 describes testing FOQUS with questions parsed from the chat logs from the Wizard of OZ experiments.

The design of question answering component takes its cues from previous ontology-based QA approaches. We focused on some of the classic implementations (NLP-Reduce (Kaufmann et al., 2007) and FREyA (Damljanovic et al., 2011), for example) that aims to be domain independent, meaning the system is not tightly coupled with domain specific aspects. This would allow any type of ontology to be queried using natural language questions. Our goal was to have a portable system that is relatively light-weight, simple, and would not require any modification of the ontology or the QA component. Overall, our QA system borrows ideas used in previous studies to create a workable system, hence the “Frankenstein” name. The figure in Appendix D outlines the system process.

FOQUS begins with importing an ontology knowledge base (Figure 4.15) where Object Property Assertions, Data Property Assertions, and Class Assertion-based axioms are extracted. These axioms are generally the core domain knowledge from which user questions will query. Object Property Assertions are basic instance-level triples and Data Property Assertions are instance-level triples attributing data to the entity-level instances. Class Assertions are domain Tbox axioms. The delineation of these types of axioms would later serve in ranking and selection of answers to be discussed later. In Figure 4.15, after the specific axioms are extracted, the domain (i.e. subject), property (i.e. predicate), and range (i.e. object) are parsed and identified. This would later serve as tuples used for comparisons. FOQUS analyzes the user’s question by extracting the noun phrases and verb phrases, and identifying the question type (Figure 4.16). The extraction of noun phrases and verb phrases are preformed by

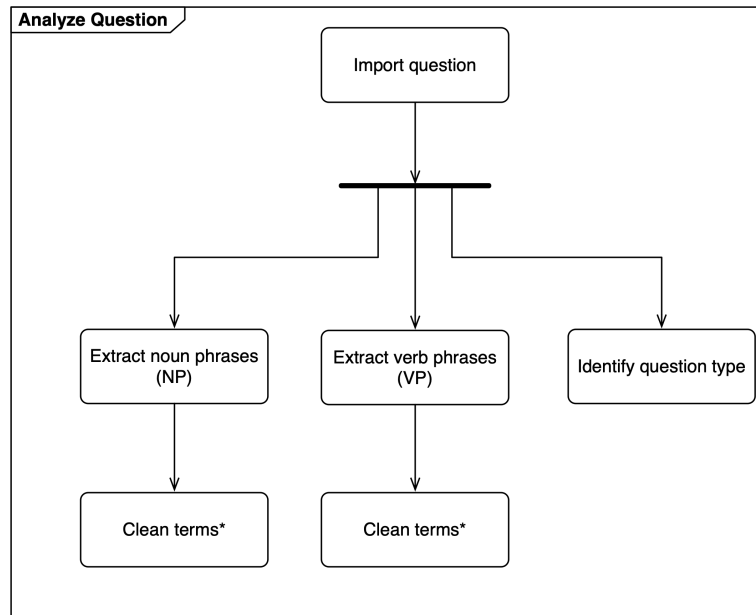


**Figure 4.15:** Extraction process from ontology for FOQUS (Appendix D)

Stanford Core NLP. The question type identification is based on NLP-Reduce’s classification which is rooted in looking at a series of keywords. For example, if the question contains “how many” or “the number of”, the question is classified as COUNT-based question. The classification has six categories - UNKNOWN, ALL (list all results), COUNT (count the results and give back the total), MAX (requesting maximum value), MIN (requesting the minimal value), and NUMERIC.

Figure 4.16 also describes FOQUS step in cleaning the terms from the noun and verb phrases. This would include removing special characters like underscores, removing duplicate words, removing common words (based on Oxford’s top 100 words), and normalizing the word to their root using MorphaStemmer.

After extracting the axiom assertions from the ontology and the question data, FOQUS computes the similarity scoring to determine what triple among axiom assertions are a probable answer for



**Figure 4.16:** Question analysis for FOQUS (Appendix D)

the question. Figure 4.17 describes the method for scoring. We utilized two methods for similarity – vector-based approach using Numberbatch as the vector model, and string-based similarity. For the later, we used the MongeElkan method, which is the method that FREyA uses for their similarity matching. By default, the Simmetrics library uses the Smith-Waterman-Gotoh for MongeElkan <sup>§</sup>, instead of Jaro-Winkler, as its root metric.

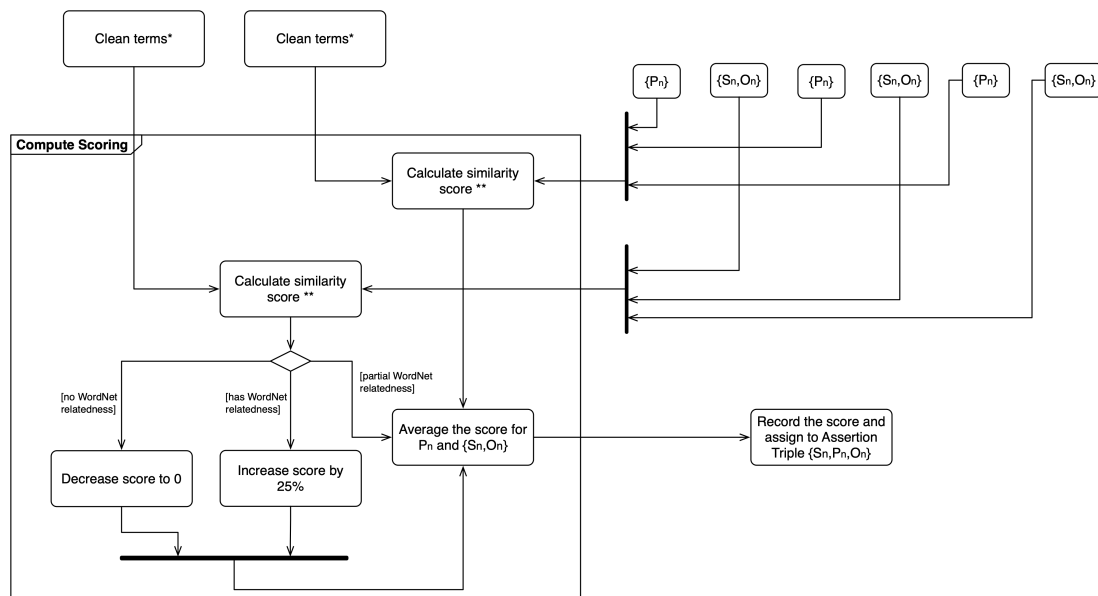
The process for determining similarity compares the predicate from triple with the verb phrase from the question. Similarly, FOQUS uses entities (subject and object) from the triple and compares it with the noun phrases from the question. In certain cases, the verb phrase was non-existent in the question so any comparison with the predicate of a triple would be ignored. All triples are sourced from the Object Assertions, Data Assertions, and the Class Assertions.

Some random, initial experiments with a sample of questions provided better scoring using WordNet

<sup>§</sup><https://github.com/Simmetrics/simmetrics/blob/59dc148f402da6a8a82ad8604a64fa33d1f70460/simmetrics-core/src/main/java/org/simmetrics/metrics/StringMetrics.java>



to enhance the resulting score. Using extJWNL, if two terms were deemed as synonymous within WordNet (using graph depth of 3), the score is increased by 25%. If there are no synonyms, hypernyms, and hyponyms relationships between the terms, the score (even if there was some similarity indicated by the two methods), would be decreased to 0. Otherwise, the scores are left as is. Lastly, the average between predicate and entity scores are recorded for the axiom triple.

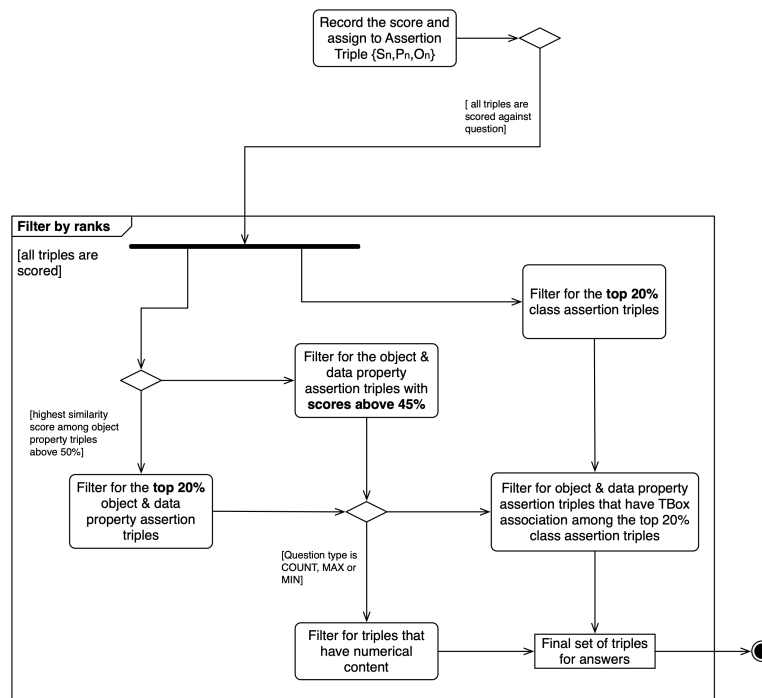


**Figure 4.17:** Similarity scoring between question data and ontology data for FOQUS (Appendix D)

The next step for FOQUS is filtering for the answer triple using the recorded scores (Figure 4.18). After all of the Object Property, Data Property, and Class Assertion triples are scored against the entities of the question, FOQUS captures the highest similarity score of the Object Property Assertion triple. If that top similarity score is above 50%, the top 20% of the Object Property and Data Property Assertions are captured. If this condition was not met, FOQUS defaults to filtering for the Object Property and Data Property Assertions above 45%. FOQUS utilizes the question type to determine additional filtering so if a question was identified as COUNT, MAX, or MIN, the system looks for triples among the selected Object Property and Data Property that have numerical content. For

example, if the triple contained “one” or “i” in its label, that triple would be selected.

If the question type were neither of the mentioned, FOQUS uses the top 20% scored of the Class Assertion triples for further selection. Using the URI for the triple’s domain, property, and range, FOQUS harnesses OWL-API and the reasoner (HermiT) to query for their respective TBox assertion. If that assertion was among the 20% of the Class Assertion triples, the Object or Data Assertion triple was selected. For example, the Object Assertion triple, *throat\_cancer*  $\rightarrow$  *affects*  $\rightarrow$  *males*, is instantiated from  $\{Disease, Target\} \rightarrow \{affects\} \rightarrow \{Males, People\}$  (if we were to include the non-direct classes). If *Disease*  $\rightarrow$  *affects*  $\rightarrow$  *People* is among the top 20% from the Class Assertion triples, then *throat\_cancer*  $\rightarrow$  *affects*  $\rightarrow$  *males* is selected.



**Figure 4.18:** Selection of answer triples based on ranks for FOQUS (Appendix D)

**ANSWER COMPOSITION FOR TRIPLES** The last process involves composing a response from the final selected triples. The assumption is that the final set of triples provided answers that are within the range of what is asked by the user. However, the issue arises on how to best present the answer that appear natural language-like and how to order the triples in a coherent manner. The following is the method used to produce a response from a set of triples using a combination of the subject, predicate, and object tuples and the similarity scoring.

**Definition 4.2.1 (Object Assertion Triple)** *An instance triple  $t$  with subject  $s$ , predicate  $p$ , object  $o$  that is an object assertion has a class assertion triple association  $T$  with subject  $S$ , predicate  $P$ , object  $O$ .*

$$t_o := \{s, p, o, \}, T := \{S, P, O\}, t_o \approx T \text{ where } s \approx S, p \approx P, o \approx O$$

**Definition 4.2.2 (Compounding Triples)** *Given an object assertion triple  $t_{oa}$  with its associated class assertion triple  $T$ , and another object assertion triple  $t_{ob}$ , if their associated class assertion triples are equal but their instance level subject  $s_a, s_b$  are dissimilar, we generate a natural language statement  $nl_g$  by compounding the triples.*

$$\begin{aligned} t_{oa} \approx T_a, t_{ob} \approx T_b \text{ where } t_{oa} &= \{s_a, p_a, o_a\}, \\ T_a &= \{S_a, P_a, O_a\}, t_{ob} = \{s_b, p_b, o_b\}, T_b = \{S_b, P_b, O_b\}. \\ \text{If } T_a &= T_b \text{ and } s_a \neq s_b, \\ \text{then } nl_g &= s_a p_a o_a \text{ "and" } s_b p_b o_b \end{aligned}$$

**Definition 4.2.3 (Aggregate Triples)** *Given an object assertion triple  $t_{oa}$  with its associated class assertion triple  $T$ , and another object assertion triple  $t_{ob}$ , if their associated class assertion triples are equal but their instance level subjects  $s_a, s_b$  are similar, we generate a natural language statement  $nl_g$*

by aggregating the objects.

$$\begin{aligned}
 & t_{0a} \approx T_a, t_{0b} \approx T_b \text{ where } t_{0a} = \{s_a, p_a, o_a\}, \\
 & T_a = \{S_a, P_a, O_a\}, t_{0b} = \{s_b, p_b, o_b\}, T_b = \{S_b, P_b, O_b\}. \\
 & \text{If } T_a = T_b \text{ and } s_a = s_b, p_a = p_b, \\
 & \text{then } nlg = s_a p_a o_a \text{ "and" } o_b
 \end{aligned}$$

Definition 4.2.3.1 (Aggregate Triples) *Given an object assertion triple  $t_{0a}$  with its associated class assertion triple  $T$ , and another object assertion triple  $t_{0b}$ , if their associated class assertion triples'  $S_a, S_b$  are equal and their instance level subjects  $s_a, s_b$  are similar, we generate a natural language statement  $nlg$  by aggregating the objects.*

$$\begin{aligned}
 & t_{0a} \approx T_a, t_{0b} \approx T_b \text{ where } t_{0a} = \{s_a, p_a, o_a\}, \\
 & \text{If } S_a = S_b, P_a = P_b \text{ and } s_a = s_b, \\
 & \text{then } nlg = s_a p_a o_a \text{ "and" } o_b
 \end{aligned}$$

Definition 4.2.4 (Compound Triple for Data Property Assertions) *An instance triple  $t_{da}$  with subject  $s$ , predicate  $p$ , object  $o$  that is an data assertion. If the subject  $s_a$  and predicate  $p_a$  are similar with another data assertion triple  $t_{db}$  subject  $s_b$  and predicate  $p_b$ . We generate a natural language statement  $nlg$  by compounding the triples.*

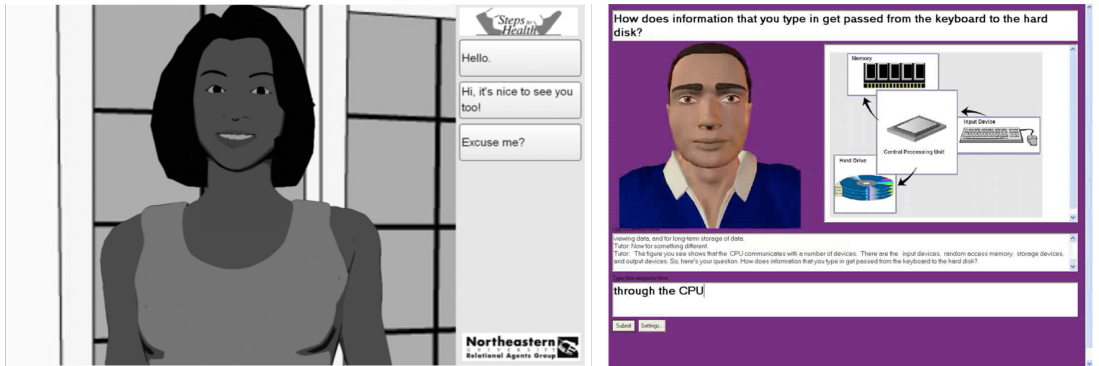
$$\text{If } s_a = s_b, p_a = p_b, \text{ then } nlg = s_a p_a o_a \text{ "and" } s_b p_b o_b$$

For the last part for the composition of the triples, we examine the highest among the final selected Object Property Assertion and the Data Property Assertion set of triples. Whichever is the highest

determines which set will precede the other set of assertion triples.

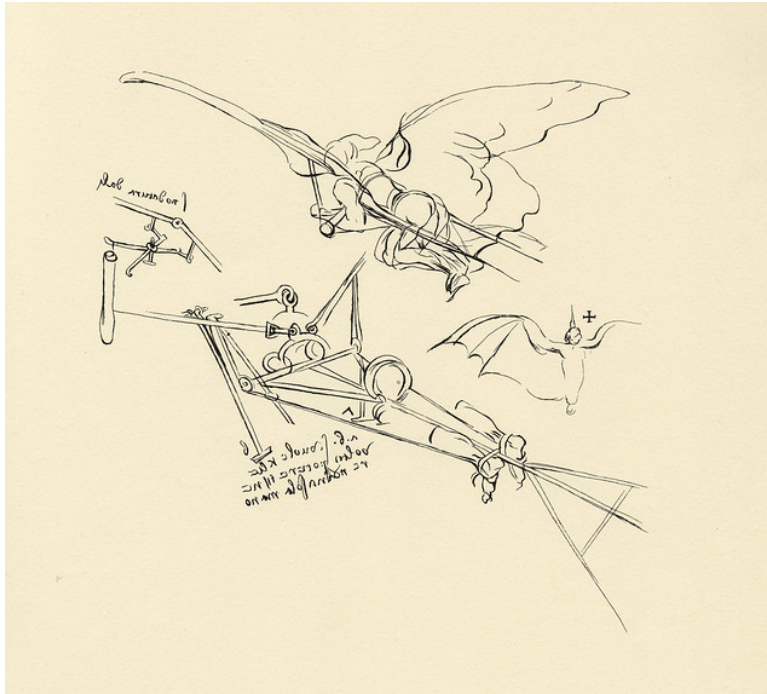
### 4.3 ONTOLOGY-DRIVEN VISUAL INTERFACE FOR AGENTS

The origination behind the visual interface design presented in this section is the fact we are solely relying on the voice and speech to communicate with the user. The data collected from users in the trials with the Wizard of OZ experiment indicated the lack of humanized voice or lack of visuals. In earlier chapters, we remarked on how speech is the natural way for us to interact with others, but there are a variety of other modalities supporting speech, like the human face that can evoke emotions. What we intend to do in this section is discuss our work in designing an innovative and unique approach to give the conversational agent a face. With some visual interface to supplement the speech we could augment the stoic-like utterances to improve the “humanization” of the conversational agent.



**Figure 4.19:** Examples of virtual agents from (Bickmore et al., 2010) on the left and on the right is AutoTutor (licensed under the CC BY-SA 3.0. <https://commons.wikimedia.org/wiki/File:AutoTutor.png>).

We break from tradition in the large body of research on virtual agents. Our reasoning is that there is difficulty in attempting to naturally mimic the human face, not to mention the time and resources to work towards a natural human face would be consuming. We attempted to experiment with an idea of using abstract visualization motifs that is acceptable for machines.



**Figure 4.20:** Leonardo da Vinci: Diagram of a proposed flying machine (1789). Licensed under CC BY-SA 2.0.  
<https://flic.kr/p/f35VPL>

For years, innovators have looked to the natural world for inspiration. Da Vinci's flying machine, a conceptual precursor to the airplane, used wings to fly that are similar to avian creatures. Yet, modern flying vehicles do not have flapping wings, either opting for static mechanical wings or rotating blades for flight. Similar to an abstract painting in a museum that can evoke or communicate a nonrepresentational idea or emotion, we experimented with the use of abstract composition of geometric features - lines, colors, shapes, to form some visual motif to provide an emotional dimension to possibly augment a conversational agent for health communication.



**Figure 4.21:** Jackson Pollock from Museum of Modern Art NYC. By cormac70 licensed under CC BY-NC-ND 2.0.<https://flic.kr/p/pZSJ1>

#### 4.3.1 VISUALIZED EMOTION ONTOLOGY

Emotions can play an integral part for health consumers in their decision making, actions and thoughts (McColl-Kennedy et al., 2017; Lerner et al., 2015). For example the emotional valence of a health consumer could determine their involvement in their personal care with their provider (Gallan et al., 2013). From a provider point of view, the providers' emotion can guide the emotion of the health consumer, a phenomenon known as emotional contagion (Hatfield et al., 1993). Overall, any interaction with a health consumer should factor in the emotion-related aspect. The work related to visual interface for the conversational agent offered an opportunity to utilize emotions that can serve as a substitute for a human-like face. We propose an ontology-driven approach which can help machines reason and understand the scope of human emotions so that conversational agents can better interact with users, particularly when discussing health information like the HPV vaccine. With the ontology, we can create dynamic visual interface for conversational agents while discussing health-related topics. In this section of the chapter, we represent visual compositions

of emotions as an application ontology and also create these visual compositions for each of the emotions defined in the ontology.

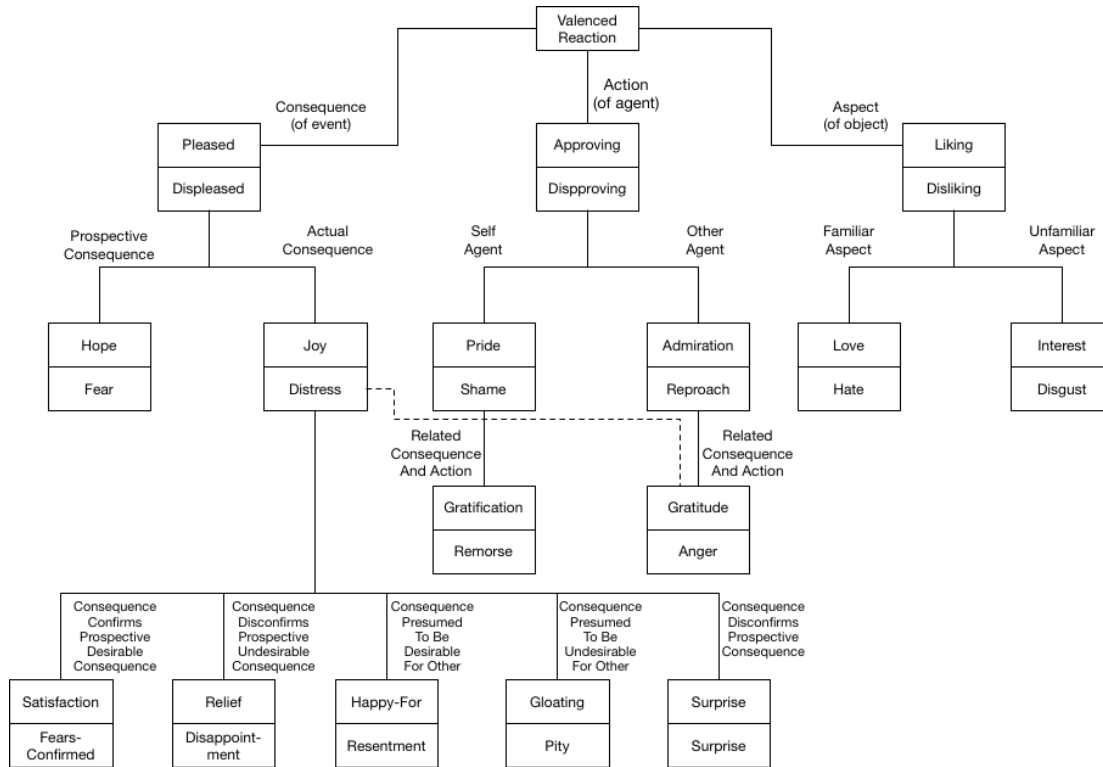
The first step in this project was to define what is an emotion or what are the emotions of a human. Many researchers in the past have proposed various models of human emotions (Ekman et al., 2013; Plutchik, 2001; Russell & Mehrabian, 1977; Lövhelm, 2012; Watson & Tellegen, 1985). We utilized a model of emotions by Ortony, Clore and Collins (OCC) that describe 22 emotions (Ortony et al., 1990; Clore & Ortony, 2013). This model was later revised in 2009 by Steunebrink and colleagues where they added interest and disgust and modified the logical structure to represent an inheritance-like structure (Steunebrink et al., 2009). Our rationale for using the OCC Model was for the following reasons:

- The revised OCC model utilizes semantics and logical structures that can be derived to be encoded in OWL2. No other study have attempted to transform this model into an ontology for distribution.
- The emotions from OCC model are defined by a collection of situational and behavioral attributes, which might be beneficial for machines to interpret assuming the machine can observe these attributes in its environment.

Aside from the revised OCC model we added the emotion of surprise from Ekman's model of emotion since all of the other Ekman's universal emotions are covered by the revised OCC. Figure 4.22 shows the final model containing 25 emotions in total. Emotion types are bifurcated as either positive or negative emotions, and then branch into three other areas - emotions that relate to the consequence of an event, emotions that relate to the actions of an agent, and emotions that involve the appraisal of an object. As an example, love is defined as a positive emotion involving linking some familiar aspect of an object, and fear is defined as a negative emotion involving some displeasure of a future consequence.

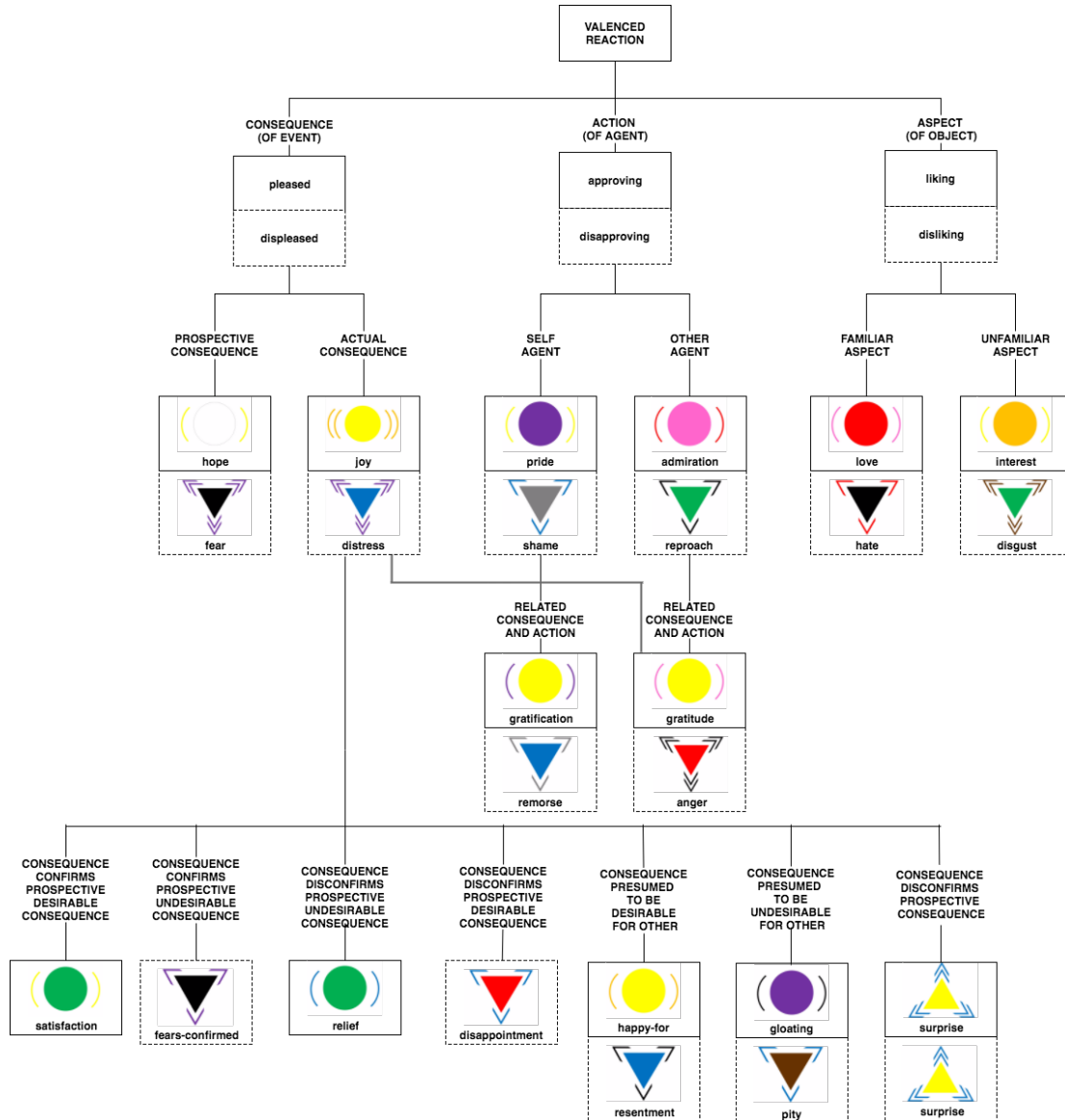
With each of the emotions, we preformed a literature review to find specific cues or visualizations associated with the emotions. From the literature, we created some composite visualizations using





**Figure 4.22:** Graphical structure of the revised OCC model that includes the emotion surprise from Ekman

shapes, lines, and colors for each emotion on the revised OCC model (See Figure 4.23)). For brevity, details of the review of the literature is discussed by [Lin et al. \(2018\)](#). Table 4.4 shows a summarization of the visualization for each of the emotion. We encoded this model using OWL2 into an distributable ontology, which we called the Visualized Emotion Ontology or VEO. Table C.1 and C.2 (Appendix C) shows the axiom definitions of the positive and negative emotions of the model.

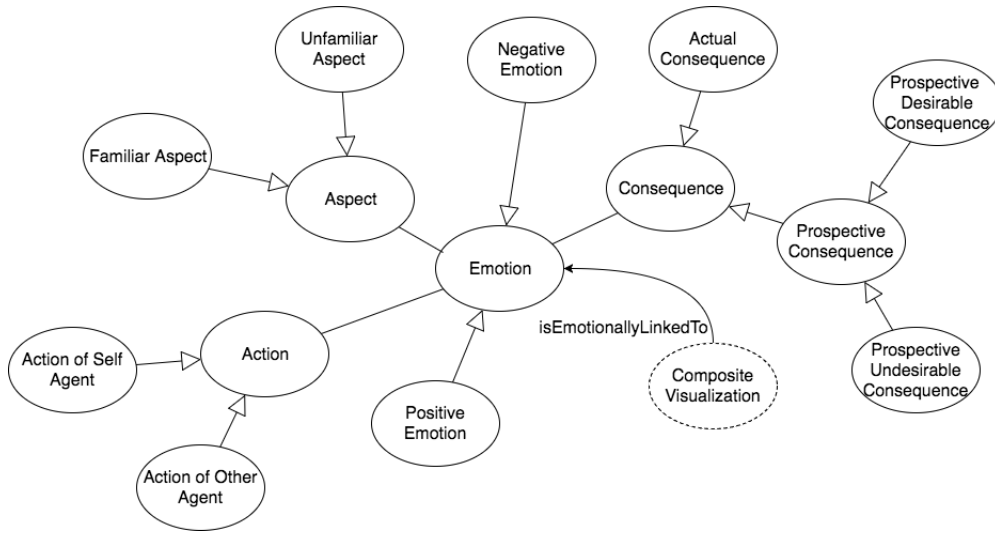


**Figure 4.23:** Graphical structure of revised OCC model (Figure 4.22) showing each emotion associated with a composite visualization.

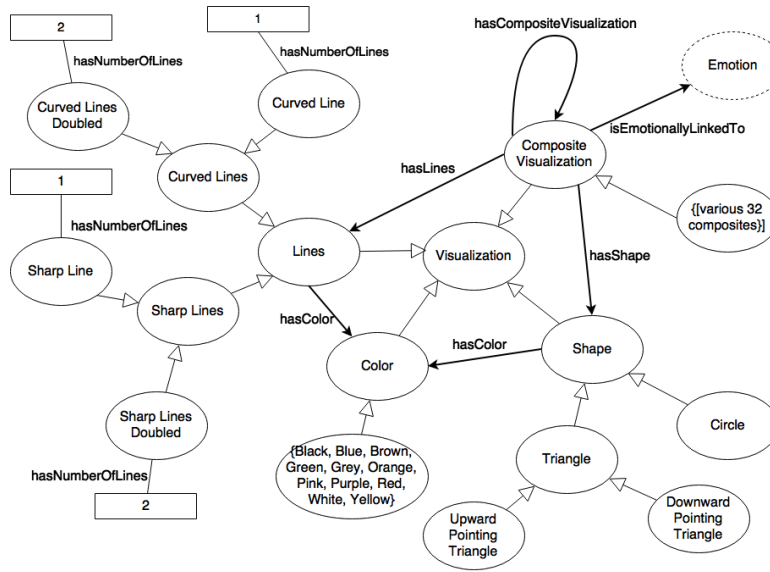
**Table 4.4:** Visualization motifs for emotions.

Emotion	Shape Attribute	Line Attribute
Admiration	Pink circle	Curved red lines, single
Anger	Red downward triangle	Curved black lines, double
Disappointment	Red downward triangle	Sharp blue lines, single
Disgust	Green downward triangle	Sharp brown lines, double
Distress	Blue downward triangle	Sharp purple lines, doubled
Fear	Black downward triangle	Sharp purple lines, doubled
Fears-confirmed	Black downward triangle	Sharp purple lines, single
Gloating	Purple circle	Curved black lines, single
Gratification	Yellow circle	Curved purple lines, single
Gratitude	Yellow circle	Curved pink lines, single
Happy-For	Yellow circle	Curved orange lines, single
Hate	Black downward triangle	Sharp red lines, single
Hope	White circle	Curved yellow lines, single
Interest	Orange circle	Curved yellow lines, single
Joy	Yellow circle	Curved orange lines, doubled
Love	Red circle	Curved pink lines, single
Pity	Brown downward triangle	Sharp blue lines, single
Pride	Purple circle	Curved yellow lines, single
Relief	Green circle	Curved blue lines, single
Remorse	Blue downward triangle	Sharp gray lines, single
Reproach	Green downward triangle	Sharp black lines, single
Resentment	Blue downward triangle	Sharp black lines, single
Satisfaction	Green circle	Curved yellow lines, single
Shame	Gray downward triangle	Sharp blue lines, single
Surprise	Yellow upward triangle	Sharp blue lines, double

We present in Figure 4.24 the basic class-level structure of the Emotion class in the Visualized Emotion Ontology. Each emotion concept, as described in Figure 4.23, is divided into a negative or positive emotion. The Emotion concept relates to a concept for the appraisal of an aspect (Aspect), as well as a Action and Consequence. Each of these branches are further specified according to the revised OCC model. Finally, each Emotion is linked to a Composite Visualization. The Composite Visualization, linked to an Emotion, has several definitions of a visualization which involves the types of Lines, Shapes, and Colors (Figure 4.25).



**Figure 4.24:** Graphical representation of the Emotion concept from VEO.



**Figure 4.25:** Graphical representation of the Visualization concept from VEO.

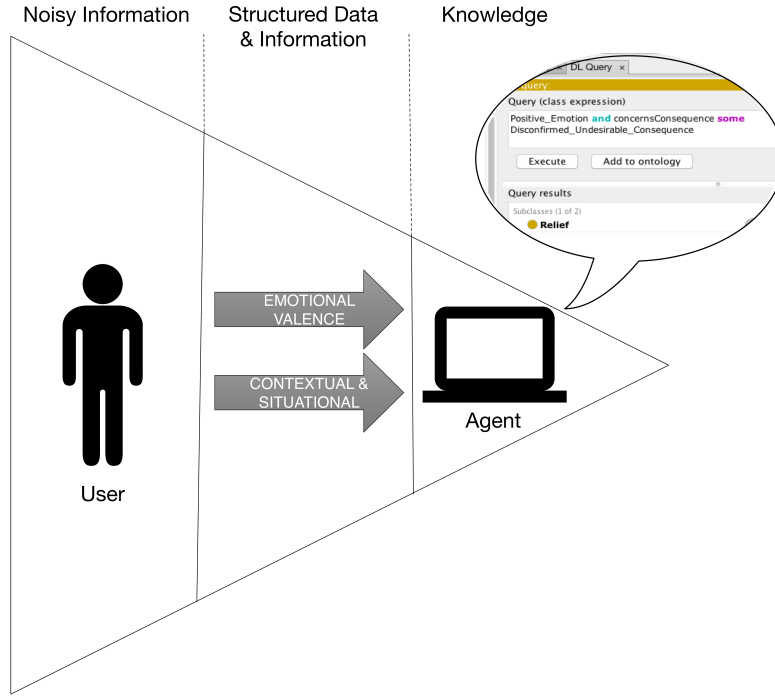
The Visualized Emotion Ontology was encoded in the OWL2 using Protégé. VEO has 126 classes, 11 data and object properties, and 25 instances (each emotion). We extended VEO from an existing emotion reference ontology called Emotion Ontology (EMO) for the sake of uniformity even

though EMO is a reference ontology and would not be an appropriate application model for interaction and visualization. Also VEO is more appropriate than EMO since EMO has responses like guilt as emotions and other behaviors that we deemed to be not emotions applicable to our needs (Lin et al., 2018).

#### 4.3.2 VEO-ENGINE

Using the Visualized Emotion Ontology that we developed, a software interface processes contextual user data to interpret emotions and provide the agent a method to express emotion by querying a linked visualization from this emotion ontology. With this software engine that we call the VEO-Engine, we can demonstrate the utility of VEO harnessed by conversational agents to express emotions using composite visualizations as its “face”.

The VEO-Engine was compiled as Java JAR executable file. It was developed using Java 8 using OWL-API, Apache Jena (Apache Software Foundation, 2009), and HermiT reasoner libraries. It uses the core of the Visualized Emotion Ontology thus links to the Basic Formal Ontology and the Emotion Ontology are not included. As stressed earlier, the ontologies discussed so far are application ontologies, and any linked knowledge base that does not offer any application purposes needed to be unlinked. This made it relatively easy to test and integrate the ontology with software. The VEO-Engine provides two basic functionalities: deducing the emotion from variables provided by the environment and retrieving a visualization to express an emotion. In Figure 4.26, the classic knowledge hierarchy (or wisdom hierarchy to some (Rowley, 2007)), but rotated to show the pathway to interpret the emotion of the user by the agent. Essentially, to reason about an emotion that is being expressed by a user, the process would involve transforming the contextual data (noisy information) and then classifying the data into structured information. The classified information can be provided to a knowledge based system (of an agent) to determine the emotion of the user. The VEO-Engine facilitates the process of consuming the input data and reasoning what the emotion



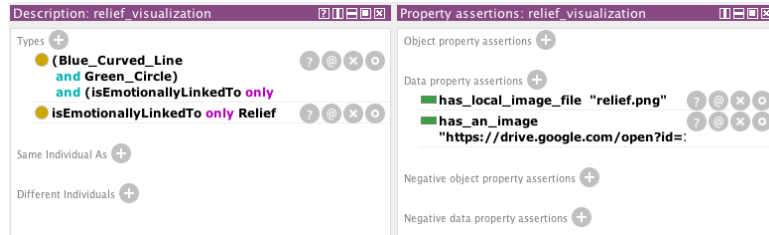
**Figure 4.26:** Mapping of the knowledge hierarchy as the process for machine interpretation of emotions.

is. The VEO-Engine utilizes the reasoning power of the VEO and HermiT reasoner library to define the emotion. The formula below is the format for the software to define emotion.

$$[positive \mid negative] \text{ and } [concept\_property_1 \dots \text{and } concept\_property_n] \quad (4.1)$$

Using our previous published example, the emotional valence is positive and the classified contextual data relates to liking something or someone that is familiar to them, the machine will reveal that the emotion is love. Love according to its definition in VEO is a *[positive]* emotion that concerns aspects of some familiar aspect (*[concept property := "concernsAspect some Familiar\_Aspect"]*).

The other functionality involves expressing emotion. Earlier we discussed how the VEO links every OCC-based emotion to a visualization. Figure 4.27 shows a screenshot of Protégé where an instance of relief visualization has a link to a web-based image and local image of the corresponding



**Figure 4.27:** Protégé screenshot of the relief visualization with a link to an image on the Web and a local image.

visualization. The system simply takes the name of the emotion that the agent wants to express and executes the SPARQL query, shown below (Listing 4.1), that retrieves the linked visualization from VEO.

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX veo: <http://sbmi.uth.tmc.edu/ontology/VEO#>

SELECT ?link WHERE
{
  ?viz veo:has_local_image_file ?link .
  ?viz a ?c .
  ?c rdf:type owl:Restriction .
  ?c owl:allValuesFrom [EMOTIONHOLDER]
}
```

**Listing 4.1:** SPARQL query for a visualization from VEO

In Chapter 5, we review the evaluation of the Visualized Emotion Ontology. Also we discuss the evaluation for the utility of VEO by evaluating the visualization through crowdsourcing. We will

also discuss the tests relating to the VEO-Engine to further demonstrate its utility.

#### 4.4 SUMMARY

To recap, this chapter presents the components of the interaction aspect of the proposed conversational agent for the HPV vaccine. The work initiated from knowledge gathered from the Wizard of OZ experiments on prospective decision makers of the HPV vaccine conversation agent - parents with children under 18 and young adults. This led to development of ontology-driven interaction tools that would be integrated in the conversational agent. One of them is the Patient Health Information Dialogue Ontology (PHIDO), an ontology that formalizes the delivery of health information using dialogue discourse modeled in OWL2. The other was the Visualized Emotion Ontology (VEO) that serves as an innovative knowledge base for interface visualizations for software agent's emotions. Both of these ontologies are supported with their respective software engines - Conversational Ontology Operator and the VEO-Engine. In Chapter 5, we present some results of the efforts of this chapter.

Lastly, we want to acknowledge and thank Rebecca Lin for her significant contribution to the Visualized Emotion Ontology that made much of the overall work possible.



*Software and cathedrals are much the same – first we  
build them, then we pray.*

Samuel T. Redwine, Jr.

# 5

## Architectural assessment and analysis of components

In this chapter, we examine the works previously discussed in Chapter 3 and Chapter 4. To recap, Chapter 3 introduces the development of domain ontologies for consumer health knowledge for the HPV vaccine. These ontologies serve as the agent's knowledge base layer from which the agent will query to power the agent's knowledge of the domain. These domain ontologies are VISO and its HPV-vaccine derivative VISO-HPV that encodes the body of knowledge found in patient-directed VIS documents. Complimenting the VISO-HPV ontology is the VAXMO ontology which anatomize misinformation of vaccines. The primary focus of this ontology is to expand the domain

scope of VISO and VISO-HPV. As mentioned in an early chapter, the extensibility of ontologies are an important benefit.

Chapter 4 discussed a couple of ontologies with a specific application-based purpose. One of the discussed ontologies was the Visualized Emotion Ontology which represented the OCC model of emotions and linked each emotion entity to an abstract visualization. The purpose of that ontology was to support affective responses of an agent – interpreting emotions and expressing emotions. To enable this we had developed the VEO-Engine that harness the VEO knowledge base. Another ontology mention in the same chapter was Patient Health Information Dialogue Ontology (PHIDO). PHIDO, inspired by our experiment with the Wizard of OZ protocol, illuminated the aspects of dialogue for patient interaction concerning HPV vaccine with a conversational agent. It describes the various utterance types that are exchanged and provided a framework to codify sequence of utterances for dialogue management. PHIDO was designed to be open to any health-related communication tasks involving an automated agent. Similar to our work with VEO, we had developed a software wrapper called Conversational Ontology Operator (COO). COO is backed by the PHIDO ontology which regulate the dialogue flow and maintains the contextual dialogue data, both of which are primary functionalities of dialogue managers. To complement COO, we also discussed the development of an automated QA subsystem, Frankenstein Ontology Question-Answering for User-Centric System (FOQUS), to respond to user questions while the COO engine is active. FOQUS utilizes an ontology as knowledge source and processes the user question to retrieve an answer from the ontology.

Our work in Chapter 3 and 4 culminated to the next step of evaluating the various components of the ontology-driven conversational agent, in order to prove the feasibility of ontology-based framework for automated agents. We start by reviewing the quality of the ontologies from the domain layer - VISO-HPV and VAXMO. Afterwards, we test the various ontology-driven software components for the application layer. Lastly, we reveal our analysis of user experience data collected

from the simulated trials of a HPV vaccine conversational agent, and present some evidence of perception and attitudes affected through exposure with the agent.

## 5.1 EVALUATION OF THE DOMAIN ONTOLOGY

We utilized a semiotic metric suite introduced by Burton-Jones and colleagues that measures an ontology based on the branches of semiotic theory (*semantic*, *syntactic*, and *pragmatic*) (Burton-Jones et al., 2005). Each of the scores in the metric range from 0 to 1, and the composite of the scores provided an overall score –  $((0.33 \cdot syntactic) + (0.33 \cdot pragmatic) + (0.33 \cdot semantic))$ . OntoKeeper is a web-based tool we have developed that calculates metrics rooted in semiotic theory - *semantic*, *pragmatic*, and *syntactic*. OntoKeeper is powered by OWL-API (Horridge & Bechhofer, 2011) and other natural language processing libraries to parse and calculate the data from the ontology. These metrics were introduced by Burton-Jones, et al. and have been used in some previous studies to evaluate ontology artifacts (Afify et al., 2017; Jianliang & Xiaowei, 2008). The benefit of this metric according to the authors, is that it is domain independent and applicable to measuring the quality of ontologies of any domain, and concise and easy to interpret and to use for evaluators (Burton-Jones et al., 2005). OntoKeeper automates the calculations of each of the metrics except for the metrics that involve external participants (i.e. subject matter expert review). The user uploads their ontology and the tool parses and extracts the meta-data needed to calculate the scores and presents them in an easy to use interface. Each of these metrics qualitatively measures the lexical quality of the concept labels (*semantics*), the domain coverage and domain applicability of the ontology (*pragmatic*), the quality of syntax for machine-readability (*syntactic*), and the community usage (*social*). For review of the semiotic evaluation scoring for ontologies see (Burton-Jones et al., 2005; Amith & Tao, 2017) for a primer. As a benchmark, we used the National Center for Biomedical Ontology (NCBO) BioPortal sample evaluation scores from our previous work (Amith & Tao,

2017).

### 5.1.1 VACCINE MISINFORMATION ONTOLOGY QUALITY ASSESSMENT

**Table 5.1:** Comparison of quality scoring derived from semiotic metric suite (Burton-Jones et al., 2005) for VAXMO and the NCBO BioPortal sample from (Amith & Tao, 2017).

\* scores and values from (Amith & Tao, 2017).

\*\* Overall score does account for social quality scores reported in (Amith & Tao, 2017).

Quality Metric	VAXMO	NCBO Sample ( $\sigma$ )*	z-score
Syntactic	0.69	0.64 (0.14)	0.36
Lawfulness	0.95	0.92 (0.16)	0.19
Richness	0.44	0.36 (0.18)	0.44
Semantic	0.94	0.88 (0.15)	0.40
Interpretability	0.91	0.88 (0.14)	0.21
Consistency	1.00	0.84 (0.40)	0.40
Clarity	0.95	0.96 (0.13)	-0.08
Comprehensiveness	<0.00	0.02 (0.07)	-0.29
Pragmatic	<0.00	0.02 (0.07)	-0.29
Overall Score	0.54	0.51 (0.07)**	0.43

We produced some initial scoring to determine an early evaluation (Table 5.1) of VAXMO's quality using our in-house web application, OntoKeeper (Amith et al., 2018b, 2019a). The *syntactic* score, which measures syntax-level assessment of the ontology (i.e. machine readability) based on any breach of syntax (*lawfulness* metric) and utilization of ontology features (*richness* metric) was 0.69, with *lawfulness* and *richness* at 0.95 and 0.44, respectively. The *semantic* score, a score that measures the term label quality of the ontology was rated at 0.94. The *semantic* score is comprised of a *consistency* score that quantifies inconsistent labeling of concepts and instances was 1.00, *clarity* that quantifies ambiguity of the term labels was 0.95, and *interpretability* that measures the ontology's term labels' meaning was 0.91.

For the *comprehensiveness* score (a component of *pragmatic* score to assess the utility of the ontology), we utilized the seed number of 1,277,993, which is the average number of classes, instances, and

properties from a sample of NCBO Ontologies in a previous study (Amith & Tao, 2017). Ideally, we would like to have identified appropriate ontologies that are comparable to VAXMO, but for initial scoring we settled on the aforementioned seed number from the previous study. *Comprehensiveness* score from the NCBO seed number provided a very low number value of less than 0.00. The *overall quality* score based on equal weighting of *syntactic* (0.69), *semantic* (0.94), and *pragmatic (comprehensiveness* at less than 0.00) was 0.54. A summary of the scores are presented in Table 5.1.

We calculated the *z-score* using the data from the NCBO BioPortal scores to attain an initial evaluation. When comparing the *syntactic* score, *z-score* yielded 0.36 indicating above-average syntactic score for VAXMO. The *z-score* for *semantic* was 0.40 also indicating above-average *semantic* score for VAXMO, and the *z-score* for *pragmatic* was -0.29 revealing below-average rating for VAXMO. Also, we calculated the *z-score* for the final *overall quality* using the average NCBO *overall score* (0.51) that does not account for the *social* metric. The *z-score* for the overall score of VAXMO was 0.43, which is above average in its overall quality compared to the NCBO sample.

We examined the *z-score* to assess the quality of VAXMO. The *syntactic* score of VAXMO appear to be of higher quality with the NCBO BioPortal sample ( $z=0.36$ ). We interpreted this to mean that the encoding of the ontology with respect to utilization of formal logic (*richness*) and minimal syntactic violations (*lawfulness*) is better than other ontologies. The *semantic* score for VAXMO was also better than the sample NCBO BioPortal ontologies ( $z=0.40$ ) with respect to minimal inconsistencies with the term labels (*consistency*), and with respect to meaningful term labels, i.e. at least one word sense (*interpretability*). However, *clarity* was slightly weaker than average ( $z=-0.08$ ), where there may have been term labels that had ambiguous meaning, i.e. above average word senses. The sample from NCBO had the benefit of larger ontologies and therefore were more comprehensive in its domain coverage than VAXMO ( $z=-0.29$ ) in regards to *comprehensiveness*. Overall, with the exception of *pragmatic (comprehensiveness)*, the Vaccine Misinformation Ontology (VAXMO) is, in its current state, a relatively respectable quality ontology based on its comparison

of *syntactic*, *semantic*, and *overall quality* scores with a sample of NCBO BioPortal ontologies. The low *pragmatic* score indicates the need for greater expansion of the ontology, and we acknowledge that VAXMO still needs some refinement and expansion. In addition, we also plan on attaining a *pragmatic* score's *accuracy* score (Burton-Jones et al., 2005) that would involve public health experts to provide a review of VAXMO's veracity which would also produce a more complete *pragmatic* score.

### 5.1.2 VISO-HPV QUALITY ASSESSMENT

**Table 5.2:** Comparison of quality scoring derived from semiotic metric suite (Burton-Jones et al., 2005) for VISO-HPV and the NCBO BioPortal sample from (Amith & Tao, 2017).

\* scores and values from (Amith & Tao, 2017).

\*\* Overall score does account for social quality scores reported in (Amith & Tao, 2017).

Quality Metric	VISO-HPV	NCBO Sample ( $\sigma$ )*	z-score
Syntactic	0.69	0.64 (0.14)	0.36
Lawfulness	1.00	0.92 (0.16)	0.50
Richness	0.38	0.36 (0.18)	0.11
Semantic	0.94	0.88 (0.15)	0.40
Interpretability	0.94	0.88 (0.14)	0.43
Consistency	1.00	0.84 (0.40)	0.40
Clarity	0.92	0.96 (0.13)	-0.31
Comprehensiveness	<0.00	0.02 (0.07)	-0.29
Pragmatic	<0.00	0.02 (0.07)	-0.29
Overall Score	0.54	0.51 (0.07)**	0.43

The ontology uses the previously developed VISO model (Amith et al., 2015) to define the conceptual class level, but also includes some refinement to accommodate complex HPV vaccine knowledge.

The Vaccine Information Statement Ontology For Human Papillomavirus (VISO-HPV) (Wang et al., 2016) contains 160 triples, 141 classes (125 subclasses), 52 properties (36 object properties and 16 data properties), and 55 individuals. Using OntoKeeper, we automatically generated the *syntactic*, *pragmatic*, and *semantic* scores and present the results on Table 5.2. The *overall quality* score based

on the three quality aspects amounted to 0.54. The *syntactic* score is 0.69 and is comprised of *lawfulness* and *richness* qualities, which are 1.00 and 0.38, respectively. *Semantic* aspect of the ontology rates at 0.94. The *interpretability*, *consistency*, and *clarity* qualities make up the *semantic* aspect, which rated, 0.94, 1.00, and 0.92, respectively. The *pragmatic* aspect score only included one factor, *comprehensiveness*, so its score is 0.00. Similar to the VAXMO, the scores comparison with the NCBO sample was similar – exceeding on all sub-scores except for *clarity* and *comprehensiveness*. The below average *comprehensiveness* is best explained because VISO-HPV is a relatively small ontology encompassing a limited of information appropriate for consumers. Most likely, VISO-HPV may have some ambiguous labels – just as VAXMO has – which explains the below average *clarity*.

### 5.1.3 VISUALIZED EMOTION ONTOLOGY (VEO) QUALITY ASSESSMENT

The VEO was encoded in the Protégé ontology authoring tool (Musen et al., 2015) in OWL2 format. The ontology contains a total of 126 classes, 11 object and data properties, and 25 instances. We scored the quality of the VEO using OntoKeeper. We compared the VEO to a sample of five cognitive ontologies (Mental State Assessment, Emotion Ontology, Mental Functioning Ontology, the Behavior Change Technique Taxonomy, and the Cognitive Atlas Ontology), which would provide us with a baseline measurement. Results of our comparison are presented in Table 5.3.

For the VEO, the *syntactic* score, a score that measures the machine-readability of the ontology, based on breaches of syntax (*lawfulness* metric) and utilization of ontology features (*richness* metric), was rated at 0.76, with *lawfulness* and *richness* at 1.00 and 0.54, respectively. The *semantic* score, a score that measures the label quality of the ontology based on the consistency of labeling of concepts and instances (*consistency* metric), the ambiguity of term labels (*clarity* metric), and the meaning of ontology term labels (*interpretability* metric), was rated at 0.97, with *consistency*, *clarity*, and *interpretability* at 1.00, 0.99, and 0.97, respectively.

The *pragmatic* score, a score that assesses the utility of the ontology based on the *comprehensiveness*

**Table 5.3:** Quality scores comparing the VEO with cognitive ontologies.

Quality Metrics	VEO	Cognitive ontologies (z-score)
Syntactic	0.76	0.58 (1.12)
Lawfulness	1.00	0.90 (0.45)
Richness	0.54	0.28 (1.68)
Semantic	0.97	0.95 (0.61)
Consistency	1.00	0.96 (0.43)
Clarity	0.99	0.97 (0.51)
Interpretability	0.97	0.97 (0.00)
Pragmatic	0.82	0.67 (0.39)
Comprehensiveness	0.82	0.67 (0.39)
Overall Quality	0.85	0.68 (0.98)

metric (i.e., domain coverage), was 0.82. The *overall quality score* based on equal weighting of *syntactic* (0.76), *semantic* (0.97), and *pragmatic* (0.82) scores was 0.85.

We calculated the *z-scores* using the data to evaluate our metrics compared to that of the sample of cognitive ontologies. The *z-scores* for the *syntactic*, *semantic*, and *pragmatic* metrics yielded 1.12, 0.61, and 0.39, respectively, indicating above-average machine-readability, linguistic quality, and domain coverage. Also, the *z-score* for the final *overall quality* was 0.98, indicating higher overall quality for the VEO than other cognitive ontologies.

Two of our collaborators (Rebecca Lin, Chen Liang) have cognitive science backgrounds. They reviewed and conferred with each other on the ontology’s veracity, and we agreed that the ontology reflected the information described in the revised OCC model.

#### 5.1.3.1 CROWDSOURCED SURVEY FOR VEO

In total, 1082 participants were surveyed through Amazon Mechanical Turk, and for each emotion-image pair, we determined the percentage of people that disagreed (1 or 2), were neutral (3), and agreed (4 or 5) that the image represented the emotion (Table 5.4).

For the majority of the emotions (17 in total –  $p < 0.001$  for 16 emotions, and  $p = 0.014$  for



**Table 5.4:** Survey results of visualization. Green highlights indicate statistically significant results.

Emotion	Disagreed	Neutral	Agreed	P-value
Admiration	24.2 %	23.1 %	52.7 %	< 0.001
Anger	6.8 %	8.1 %	85.1 %	< 0.001
Disappointment	43.5%	23.9%	32.6%	1.000
Disgust	51.7%	17.6%	30.7%	1.000
Distress	40.3%	20.9%	38.8%	0.711
Fear	18.6%	16.4%	65.0%	< 0.001
Fears-Confirmed	21.9%	18.7%	59.4%	< 0.001
Gloating	46.8%	27.1%	26.1%	1.000
Gratification	23.6%	27.7%	48.7%	< 0.001
Gratitude	19.4%	27.5%	53.1%	< 0.001
Happy-For	10.4%	13.6%	76.0%	< 0.001
Hate	10.3%	11.8%	77.9%	< 0.001
Hope	27.0%	19.5%	53.5%	< 0.001
Interest	26.6%	27.8%	45.6%	< 0.001
Joy	9.3%	11.8%	78.9%	< 0.001
Love	29.8%	18.4%	51.8%	< 0.001
Pity	46.4%	25.3%	28.3%	1.000
Pride	25.2%	24.5%	50.3%	< 0.001
Relief	20.2%	22.6%	57.2%	< 0.001
Remorse	42.8%	25.5%	31.7%	1.000
Reproach	48.0%	31.1%	20.9%	1.000
Resentment	57.9%	22.5%	19.6%	1.000
Satisfaction	21.7%	22.7%	55.6%	< 0.001
Shame	35.1%	24.0%	40.9%	0.014
Surprise	17.2%	18.8%	64.0%	< 0.001

emotion of shame), people tended to agree that our visualization matched the emotion more than they disagreed, which validates our model; these emotions included admiration, anger, fear, fears-confirmed, gratification, gratitude, happy-for, hate, hope, interest, joy, love, pride relief, satisfaction, shame, and surprise. This conclusion is based on a rigorous hypothesis testing procedure. Specifically, we assumed that the choice of each participant was distributed as a multinomial distribution with parameters  $p_1, p_2, p_3$  corresponding to the proportions of “Disagreed”, “Neutral”, and “Agreed”. We

then performed one-sided hypothesis tests to test whether the proportion of people who agreed is greater than the proportion of people who disagreed for each of the 25 emotions, i.e.  $H_0 : p_1 < p_3$  for each emotion. Bonferroni correction was applied to control the family-wise error rate at 5%. P-values are reported in Table 5.4. Significant results of higher proportion of agreed than disagreed (p-value  $< 0.001$ ) were found for 16 out of 25 emotions including all of the emotions previously stated except for shame (p-value = 0.014).

For the remaining eight emotions, more people disagreed than agreed with our visualization. However, for five of these emotions, including disappointment, disgust, gloating, pity, and remorse, more people agreed with our emotion-image pairs than they did for the incorrect emotion-image pairs. In these cases, the randomly-selected incorrect emotion-image pairs included disappointment-interest, disgust-satisfaction, gloating-gratitude, pity-admiration, and remorse-gratification. For distress, reproach, and resentment, however, more people agreed with the incorrect emotion-image pairs than they did with the correct ones; these incorrect pairs included distress-fear, reproach-resentment, and resentment-disappointment, respectively.

#### 5.1.4 PATIENT HEALTH INFORMATION DIALOGUE ONTOLOGY (PHIDO) ASSESSMENT

In a previous study, we generated scores for a BioPortal sample to serve as comparison benchmark to assess drug ontologies (Amith & Tao, 2017). We used this benchmark comparison data to compare with PHIDO's metrics to compare its quality with other ontologies. To calculate PHIDO's scores, we imported the ontology to OntoKeeper, a prototype tool we developed that facilitates the aforementioned semiotic metric suite (Amith & Tao, 2017) and present the results on Table 5.5.

PHIDO's *syntactic* score, which measure the quality of syntax language of the ontology, was 0.69.

The sub-scores for *syntactic*, *lawfulness* and *richness*, were 1.00 and 0.38, respectively. *Lawfulness* indicates any syntactic violations to OWL2 profile. The high score of 1.00 reveals no syntactic violations. *Richness* highlights the percentage amount of unique types of logical axiom ontology features. The

**Table 5.5:** Comparison of quality scoring derived from semiotic metric suite (Burton-Jones et al., 2005) for PHIDO and the NCBO BioPortal sample from (Amith & Tao, 2017).

\* scores and values from (Amith & Tao, 2017).

\*\* Overall score does account for social quality scores reported in (Amith & Tao, 2017).

Quality Metric	PHIDO	NCBO Sample ( $\sigma$ )*	z-score
Syntactic	0.69	0.64 (0.14)	0.36
Lawfulness	1.00	0.92 (0.16)	0.50
Richness	0.38	0.36 (0.18)	0.11
Semantic	0.94	0.88 (0.15)	0.40
Interpretability	0.94	0.88 (0.14)	0.43
Consistency	1.00	0.84 (0.40)	0.40
Clarity	0.92	0.96 (0.13)	0.31
Comprehensiveness	<0.00	0.02 (0.07)	-0.29
Pragmatic	<0.00	0.02 (0.07)	-0.29
Overall Score	0.54	0.51 (0.07)**	0.43

score of 0.38 revealed that PHIDO only used about a third of these features. In comparison, with the BioPortal sample, the *z-score* for the *syntactic* score was  $z=0.36$  ( $z=0.5$  for *lawfulness* and  $z=0.11$  for *richness*) indicating a better syntactic-level quality.

The *Semantic* score measures an ontology's quality of term labels. The *semantic* score for PHIDO was 0.94. The *semantic* score comprises of *interpretability*, *consistency*, and *clarity*. *Interpretability* measured at 0.94, *consistency* measured at 1.00, and *clarity* was 0.92. *Z-score* for *semantic* score rated at  $z=0.40$  ( $z=0.43$  for *interpretability*,  $z=0.40$  for *consistency*, and  $z=-0.31$  for *clarity*). While the overall *semantic* score was better, the sub-score for *clarity* was low compared to the National Center for Biomedical Ontology (NCBO) sample's *clarity* sub-score. This may indicate that the term labels have some ambiguity (i.e. term labels that has above average number of word senses).

The *pragmatic* score assessed the ontology's domain coverage and utilization. This score was limited to its sub-score of *comprehensiveness*. The other sub-scores of *pragmatic* included *relevance* and *accuracy* which required external assessment resources (domain experts). Essentially, *comprehensiveness* measured the ontology's domain coverage based on its size in comparison with the average size of a

ontology library. The *pragmatic* score (*comprehensiveness*) when rounded to nearest two digits was 0.00. *Z-score* yielded  $z=-0.29$ , below average in domain coverage than the average NCBO ontology, which indicates that the ontology may need to be further expanded (e.g., more Speech Tasks to be modeled).

The *overall score* for PHIDO is a mean value of the previously mentioned scores ( $Q = w_{q_1} * S + w_{q_2} * E + w_{q_3} * P + w_{q_4} * O$ ) to indicate general assessment of the ontology. PHIDO's *overall score* was 0.54 and when compared, the *z-score* value was  $z=0.43$ . The *overall score* indicated that PHIDO was above average quality compared to most published ontologies, however, as noted, the domain coverage and ambiguity of term labels were lacking.

## 5.2 APPLICATION OF ONTOLOGIES FOR DIALOGUE SYSTEM COMPONENTS

The application layer of the system is supported from a combination of domain and application based ontologies. For dialogue, PHIDO was designed to manage the agent's dialogue interaction using the dialogue script tested in the Wizard of OZ experiments. In these experiments users asked questions about the HPV vaccine. To support automated question answering, we utilized VISO-HPV as the core knowledge base to query information to answer patient questions. Both the automated dialogue management and question answering was facilitated by the dialogue engine that we call COO (Conversational Ontology Operator), which also furnished a subsystem for question answering. COO utilized both the PHIDO and VISO-HPV to power the communication functionality of the agent. VEO-Engine facilitated the visual interface as a means to express motive interactions with users. This engine was supported by the aforementioned VEO that encodes the model of emotions defined by the OCC model with links to abstract visualizations. The engine also utilized the ontology's reasoning capabilities to interpret emotions of the user based on situations. In the following subsections, we discuss some experimental tests to assess the functionality of these systems

which could be leveraged by a conversational agent.

### 5.2.1 PHIDO FOR MANAGING THE DIALOGUE ENGINE (COO)

Most of the dialogue interaction is primarily communicating the singular pieces of information about the HPV and HPV vaccine. We chose to focus on the core dialogue exchange which is the communication of health information to the user as our test example. Figure 5.1 has the figure that outlines the structure of this core interaction. To assess PHIDO's ability to direct the COO engine's interaction, we present the following:

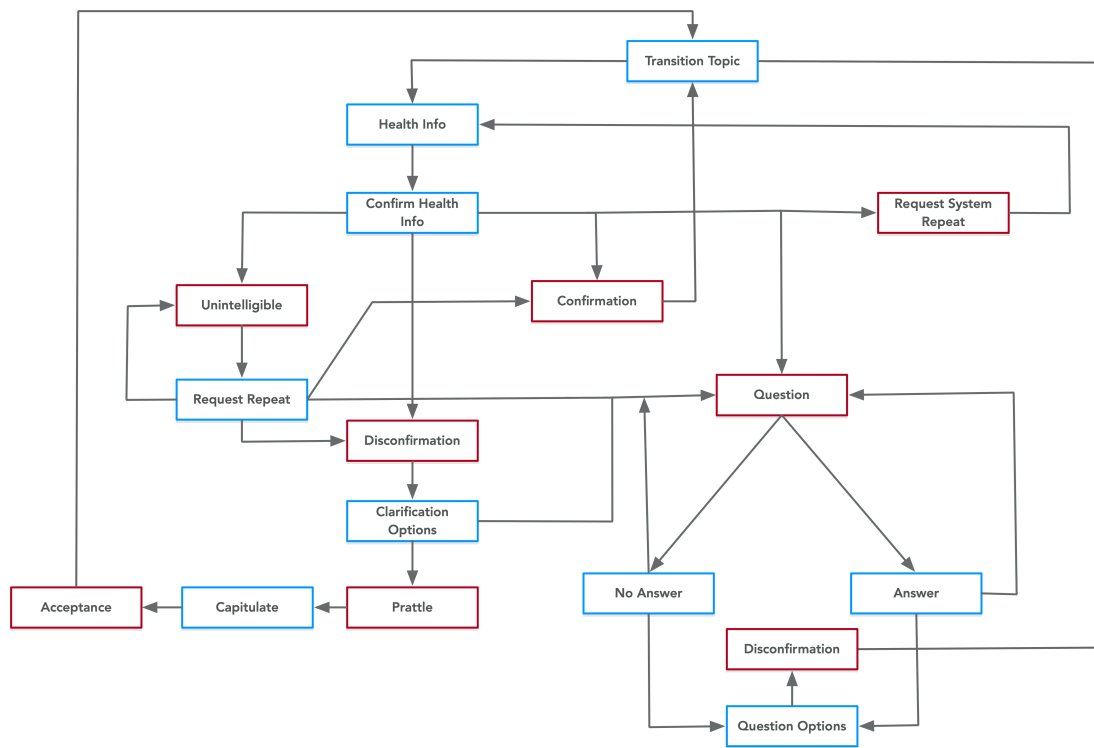
Can PHIDO direct the agent's engine (COO) to

1. Impart a piece of health information (HPV vaccine related) to the user?
2. Coordinate question answering?
3. Transition the conversation to discuss a health topic?

For the first question, we tested several use cases. One of the use cases was to assess if the system can handle the user confirming they have heard the health information. Another use case was to review the system's ability to manage if the user did not agree or did not understand the health information communicated to them. Other use cases included requesting repeating the health information, switching to question answering mode, and handling misunderstanding from the user. For the second question, we tested the engine's ability to provide the answer or no answer to the user's question and also present options for the user to ask another question. The question answering for our test cases are all simulated, but in a later section we will discuss the automated question answering that aims to be integrated as subsystem for COO. By default, all of the use cases end with transition to the next health topic to fulfill the third objective.

For our test, we created several instances in PHIDO that represented dialogue utterances, which instantiated the various Utterance classes for the Health Information Task (See Figure 4.10). Figure

5.2 shows a screenshot of Protégé with the instances populating PHIDO. In total, we had 19 instances of various Utterance classes. Each of these utterance instances were linked using a specific object property *precedes*. For convenience, the instance labels were annotated according to the classes they were derived from. This allowed us to easily assess and explain our test cases. The test cases involve sample user interaction and were enacted in a text-based console.



**Figure 5.1:** Modified version diagram of the Discuss Health Task with the Transition Topic utterance.

In the following section, we utilize an annotated-version of Figure 5.1, accompanied by the console print out to demonstrate the operation of the engine. Figure 5.1 is an adaption of Figure 4.10 with the Utterance class Transition Topic to address the third competency question relating to transitioning to the next health topic.

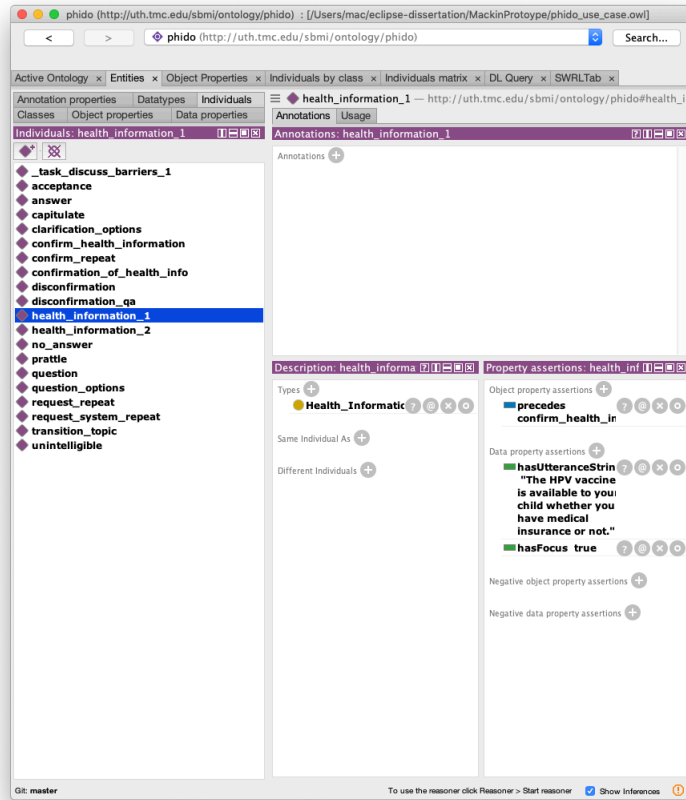


Figure 5.2: Screenshot of Protégé with PHIDO's instances for testing.

### 5.2.1.1 RESULTS FOR THE CONVERSATIONAL ONTOLOGY OPERATOR

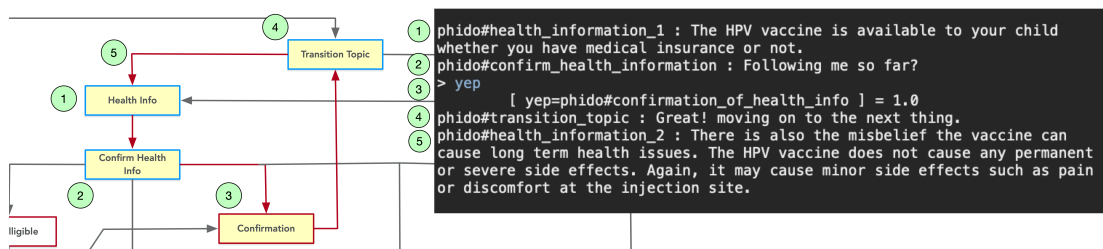
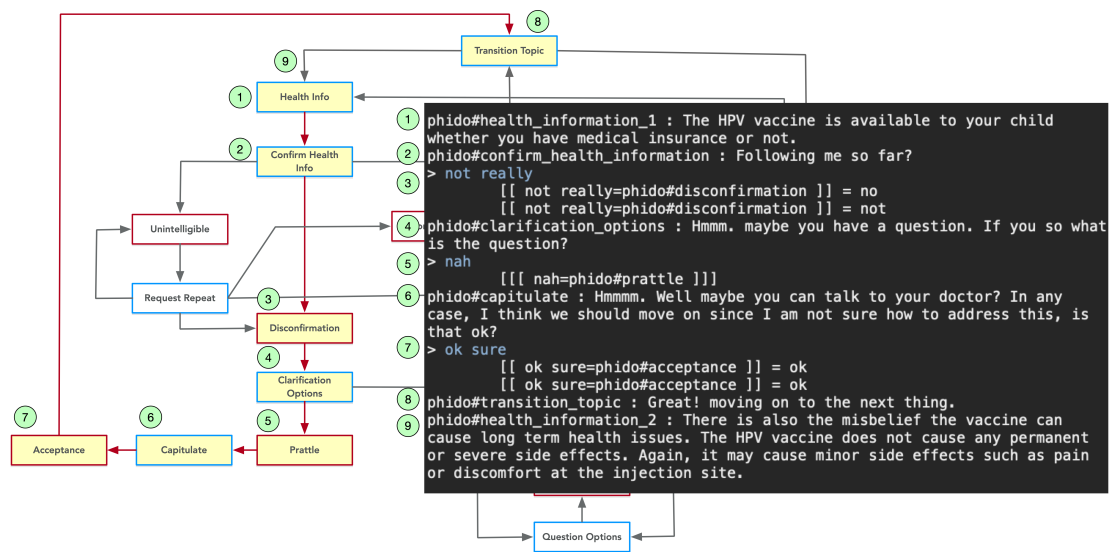


Figure 5.3: Dialogue interaction showing confirming health information. Red arrows indicate the path, and yellow box is the Utterance utilized in the result. See Figure 5.1 for a complete view of the flow diagram.



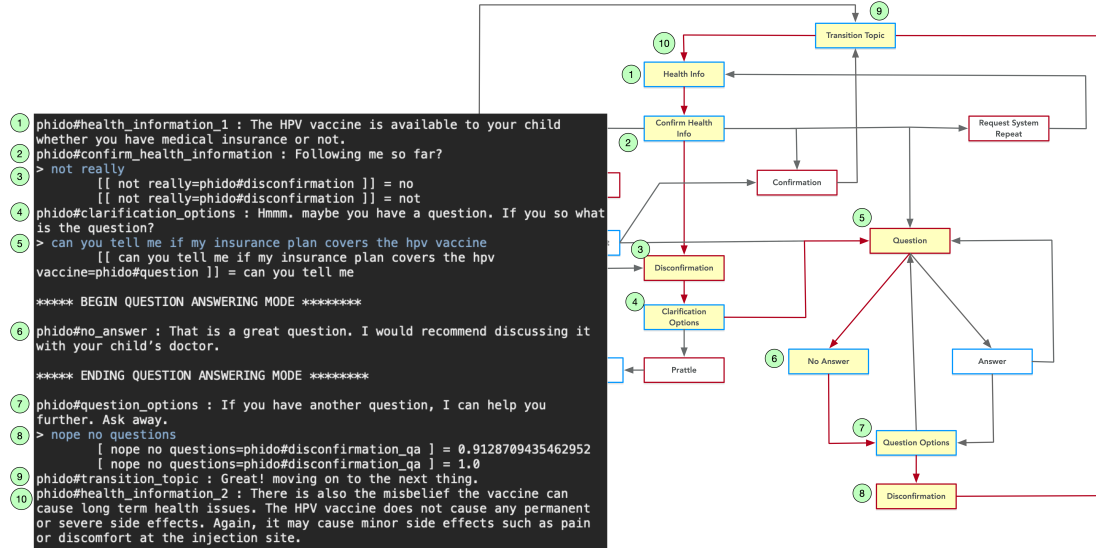
**Figure 5.4:** Dialogue interaction showing disconfirming health information. Red arrows indicate the path, and yellow box is the Utterance utilized in the result. See Figure 5.1 for a complete view of the flow diagram.

For the test case revolving around the user indicating that he/she understands the health information communicated to them, Figure 5.3 shows text console demonstrating the test case. For this case, the COO engine tells the user the HPV vaccine is available irrespective of their insurance status and then follows up with the agent asking whether the user confirms this information. In this assessment we responded with a “yes” and the engine identifies that it is a confirmation. The engine declares the dialogue continues to the next health information. The contrast to the previous use case is if the user misunderstands or has some contentious notion of the information provided. Figure 5.4 outlines the test case with the a simulated user saying “not really” in response to the health information uttered. The engine identifies the utterance as Disconfirmation and directs the agent to inquire if they have a question. The response is negative (e.g. “nah”) which the agent understands as prattle. The engine directs the agent to ask the user to move on to the next topic and that it is best to ask their health care provider if there is an issue.

If the user wants the agent to repeat the engine can facilitate repeating the same health information



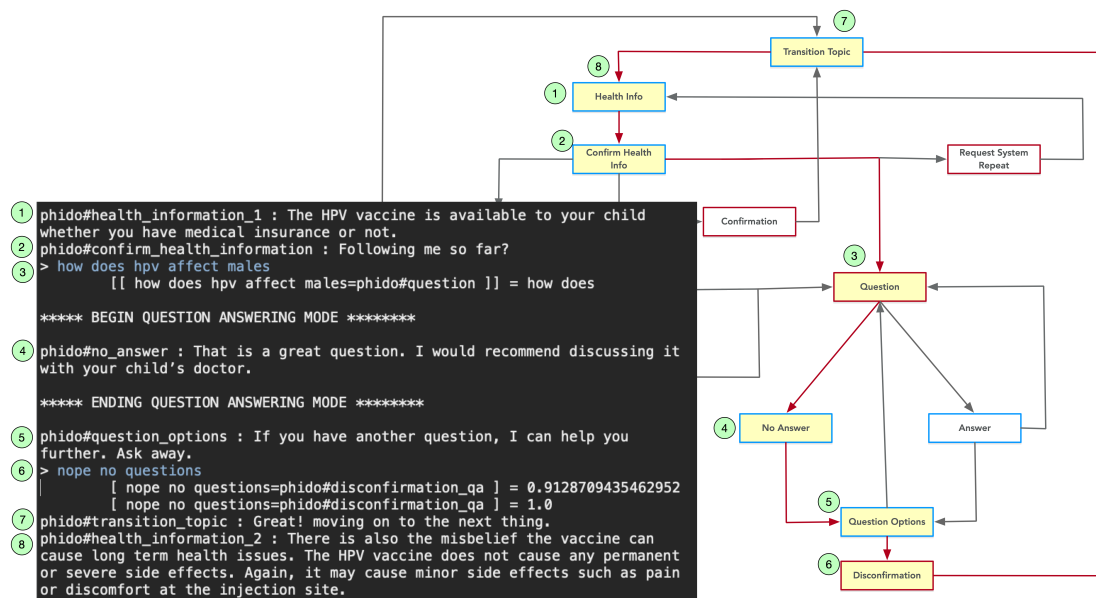




**Figure 5.7:** Dialogue interaction showing the transition from health information exchange to question answering mode (simulated). Red arrows indicate the path, and yellow box is the Utterance utilized in the result. See Figure 5.1 for a complete view of the flow diagram.

COO, with direction from the PHIDO ontology, can handle situations where there may be a misunderstanding between the user and machine. Figure 5.6 shows an example, albeit a humorous situation, that highlights the engine's ability to handle a use case where confusion may happen. Figure 5.6 has a series of exchanges from the user that was identified as Unintelligible class to rectify any confusion and then segueing to the next health topic to discuss.

Figure 5.7 illustrates the test case for one of the ways the engine can switch to question answering mode (to be facilitated by FOQUS). In this case, not really is discerned as a Disconfirmation utterance and the COO engine directs the agent to ask if the agent's user has a question. The question is provided and successfully identified as a Question utterance type, which directs COO to switch to question answering mode (simulated for test cases). The simulated question answering system responds (the agent does not have an answer). The utterance "nope no question" detected as Disconfirmation utterance type signals the COO engine to continue. Figure 5.7 displays the details of the exchange for this use case.

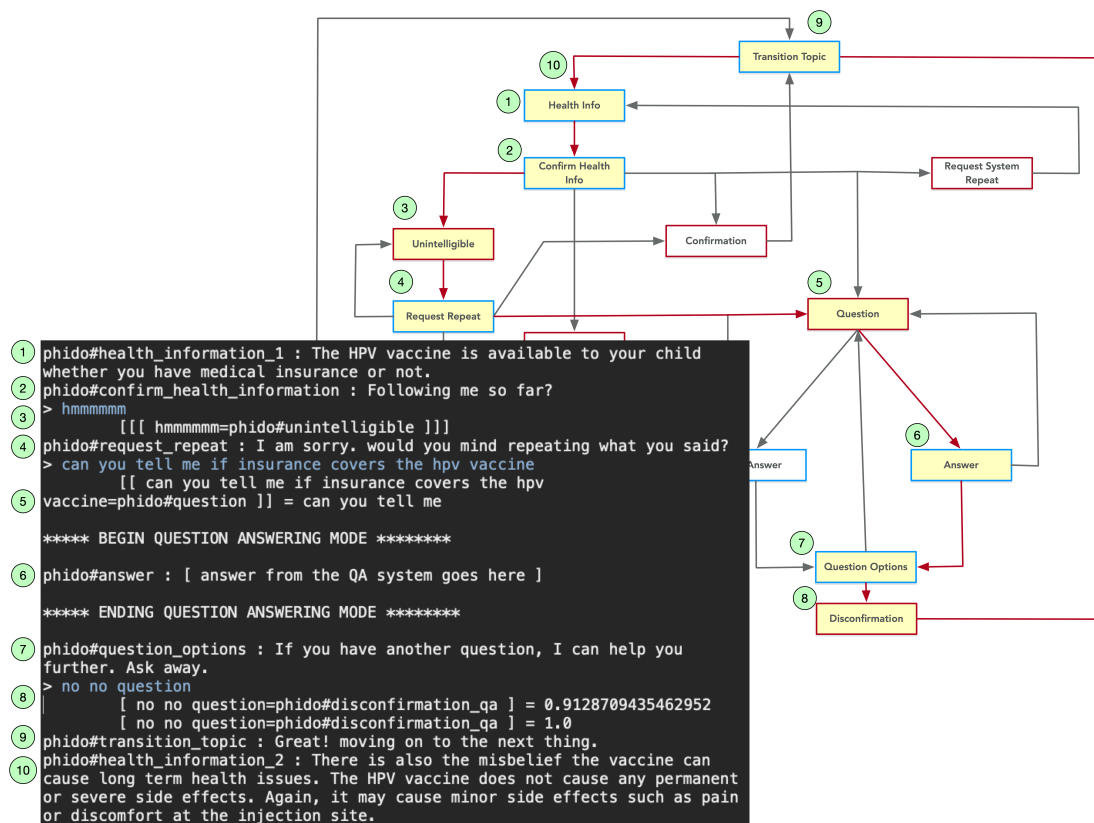


**Figure 5.8:** An alternate dialogue interaction showing the transition from health information exchange to question answering mode (simulated). Red arrows indicate the path, and yellow box is the Utterance utilized in the result. See Figure 5.1 for a complete view of the flow diagram.

Another way in the dialogue interaction to direct the agent to question answering mode is demonstrated in Figure 5.8. The use case is similar to the previous, with the difference of the user asking a question when the agent inquires if the user confirms the information communicated to them.

Figures 5.9 and 5.10 show a similar dialogue interaction for answering a question, one is where the agent has a response to the question and the other is where the agent has no response to the question. The question regarding whether insurance is covered by the HPV vaccine (i.e. “can you tell me if insurance covers the hpv vaccine”) is recognized as a *Question* type utterance. This directs the system to switch to question answering mode and the simulated question answering gives either an answer or no answer. Afterwards, the COO engine directs the agent to continue with the next health information content. Both Figures 5.9 and 5.10 contain details of the exchanges for the use cases.

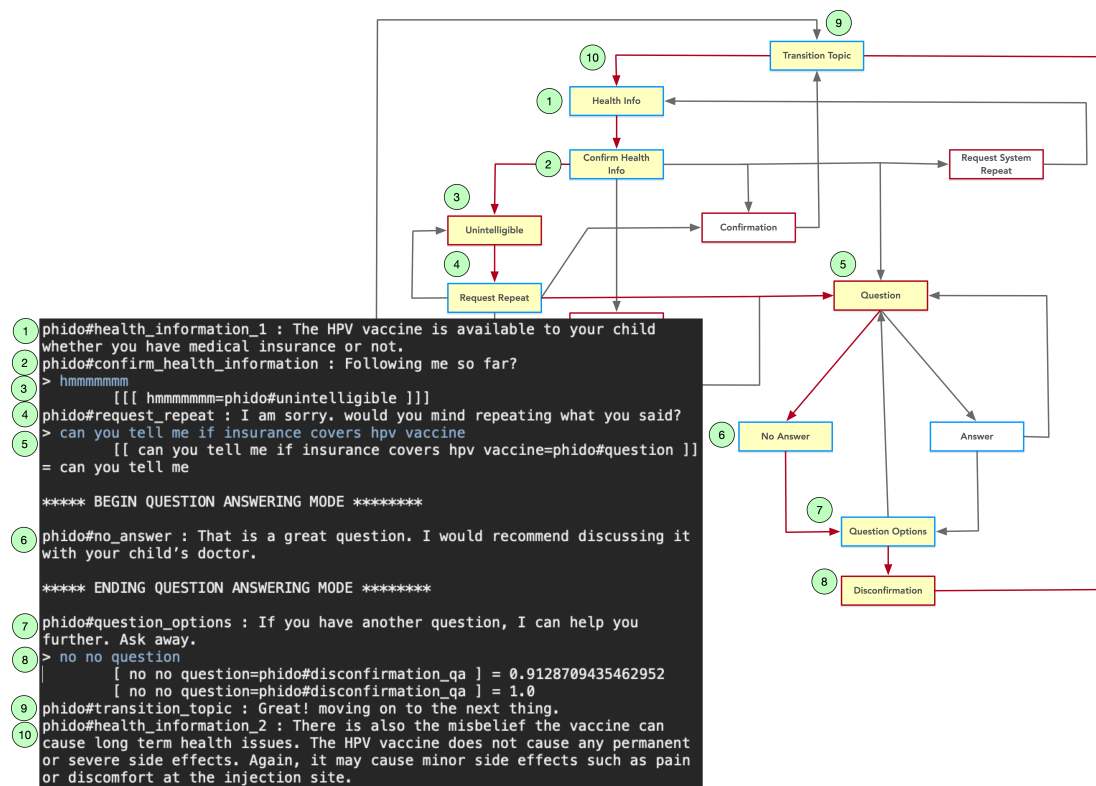
Within the question answering interaction, COO handles situations where the user may ask multiple



**Figure 5.9:** Dialogue interaction showing a question answered. Red arrows indicate the path, and yellow box is the Utterance utilized in the result. See Figure 5.1 for a complete view of the flow diagram.

questions. Figure 5.11 illustrates this use case starting from utterances that signals the COO engine to switch to the question answering subsystem. The engine facilitates the interaction for the first question (“can you tell me if insurance plans cover vaccination”) and second question (“how does hpv affect males”), then segues to next health topic. Details of the sequence of the interaction are shown in Figure 5.11.

In all of the above-mentioned use cases, by default, an instance of the next health information (health\_information\_2) is added to demonstrate COO’s movement from one Speech Task to another. In the examples provided the agent transition from the Discuss Health Task (for expressing that HPV vaccine is available regardless of insurance status) to an utterance of Health Information



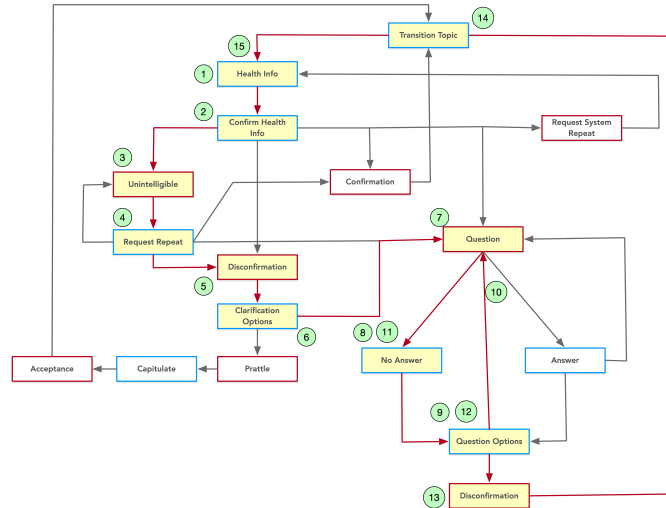
**Figure 5.10:** Dialogue interaction showing a question with no answer. Red arrows indicate the path, and yellow box is the Utterance utilized in the result. See Figure 5.1 for a complete view of the flow diagram.

class (for expressing there is a misbelief of long term effects of the HPV vaccine) of another Discuss Health Task.

```

1 phido#health_information_1 : The HPV vaccine is available to your child whether you have medical insurance or not.
2 phido#confirm_health_information : Following me so far?
3 > kinda
4 [[ kinda=phido#unintelligible ]]
5 phido#request_repeat : I am sorry. would you mind repeating what you said?
6 > nothing
7 [ nothing=phido#disconfirmation ] = 1.0
8 phido#clarification_options : Hmm. maybe you have a question. If you so what is the question?
9 > can you tell me if insurance plans cover vaccination
10 [[ can you tell me if insurance plans cover vaccination=phido#question ]] = can you tell me
11 ***** BEGIN QUESTION ANSWERING MODE *****
12 phido#no_answer : That is a great question. I would recommend discussing it with your child's doctor.
13 ***** ENDING QUESTION ANSWERING MODE *****
14 phido#question_options : If you have another question, I can help you further. Ask away.
15 > how does hpv affect males
16 [[ how does hpv affect males=phido#question ]] = how does
17 ***** BEGIN QUESTION ANSWERING MODE *****
18 phido#no_answer : That is a great question. I would recommend discussing it with your child's doctor.
19 ***** ENDING QUESTION ANSWERING MODE *****
20 phido#question_options : If you have another question, I can help you further. Ask away.
21 > nope no questions
22 [ nope no questions=phido#disconfirmation_ga ] = 0.9128709435462952
23 [ nope no questions=phido#disconfirmation_ga ] = 1.0
24 phido#transition_topic : Great! moving on to the next thing.
25 phido#health_information_2 : There is also the misbelief the vaccine can cause long term health issues. The HPV
26 vaccine does not cause any permanent or severe side effects. Again, it may cause minor side effects such as pain or
27 discomfort at the injection site.

```



**Figure 5.11:** Dialogue interaction providing the user the option to ask another question. Red arrows indicate the path, and yellow box is the Utterance utilized in the result. See Figure 5.1 for a complete view of the flow diagram.

## 5.2.2 QUESTION ANSWERING (FOQUS) FOR VISO-HPV

In Chapter 4, we outlined the subsystem that analyzes user questions and produces an answer from an ontology. This subsystem was designed to support the PHIDO-driven COO engine that manages the dialogue for HPV vaccine counseling. The question answering service is called FOQUS

and utilizes the VISO-HPV to provide a knowledge base of vaccine information for patients to query.

To test FOQUS, we used questions asked during our simulated experiment with participants (Amith et al., 2019b). In total, we collected 53 questions that range from age appropriateness for the vaccine, gender-related questions, cost, etc. Some of the questions may have been mis-transcribed from speech recognition, yet we kept them as is to imitate how the live system would process the question. Because of the possibility of mis-recognition of the utterances, FOQUS relies on the salient terms of the question (noun and verb phrases) to retrieve an answer. FOQUS provides two variants, one that employs vector similarity and the other string similarity matching. Both of these were tested against the 53 questions. Each of these questions was imported into the FOQUS system and FOQUS outputted answers for each of the questions.

We enlisted the help of four evaluators and asked them to evaluate the question and answer pairs based on two criteria: the acceptability of the answer for the questions (on a 5 point Likert scale) and whether the answer matches the question (2=yes, 1=partial, 0=no). The first criterion was devised to help us understand if the presentation and composition of the question from triples. The second criterion question helped us to determine if system can answer the question with some degree of relevancy. We calculated Cohen Kappa's inter-rater reliability Cohen (1960) for both of these questions to determine the effectiveness of FOQUS.

#### 5.2.2.1 RESULTS FOR FOQUS

We compiled the assessments for each of the questions from our evaluators. For the criterion regarding the acceptability of the answer, we normalized the ratings for degree of acceptability (5 and 4) to 1, neutral (3) to 0, and degrees of unacceptability (2 and 1) to -1. For criterion addressing whether the answer responded to the question correctly, both answered (2) and partially answered (1) were recoded as 1 and unanswered (0) coded as 0. In addition, we also tallied the non-normalized agreement

(conservative) to further assess the performance of the question answering system. Kappa's inter-rater agreement was calculated on these recoded values among the four evaluators. In Table 5.6, we present agreement results for FOQUS.

Table 5.7 presents the accuracy of FOQUS, along with the percentage of acceptability for natural language composition of the answer. Similar to above, we calculated accuracy of the question responses by coding the partially answered and completely answered as 1, and 0 coded for not answering the question. We also present the accuracy for completely answered as 1, and coded partially answered and not answered as 0. Presentation of the answer was coded as 1 for degrees of acceptability, and neutral and the degrees of unacceptability to 0.

**Table 5.6:** Agreement ratings for the question answering component.

FOQUS config	Acceptable answer	Perceived correctness (conservative)	Perceived correctness
vector variant	0.55	0.59	0.80
string variant	0.64	0.66	0.82

**Table 5.7:** Accuracy of the question answering component.

	Vector variant	String variant
Response answered the question		
<i>combined with partial</i>	0.72	0.70
<i>without partial (exact)</i>	0.54	0.50
Acceptable presentation of answer	0.50	0.49

Based on the agreement for acceptability, the semantic vector variant for FOQUS rated at 0.55. While the string-based variant rated at 0.64. For the agreement of whether the answer addressed the question, the vector-based variant rated at 0.80, and the string-based configuration rated at 0.82. The raw conservative agreement was 0.59 and 0.66 for vector and sting variants, respectively. FOQUS' vector-based variant appear to preform slightly better for answer accuracy on both exact



(0.54 to 0.50) and calculations that include partially answered (0.72 to 0.70). When considering the agreement from Table 5.6 where the string variant of FOQUS has slightly more agreement from evaluators, the better accuracy may not be conclusive. The same can be said for the presentation of the answer where the vector-based variant of FOQUS was slightly better (0.50 to 0.49).

### 5.2.3 VEO-ENGINE

The objective of this study was to show that, for small devices, we could use an emotion ontology to reason and query emotions. This study could further our work in developing conversational agents that include emotions in interactions with humans. Also, this may further interest in using ontologies and the semantic web to help machines express and interpret emotions with humans users.

To support our objective, we performed the following:

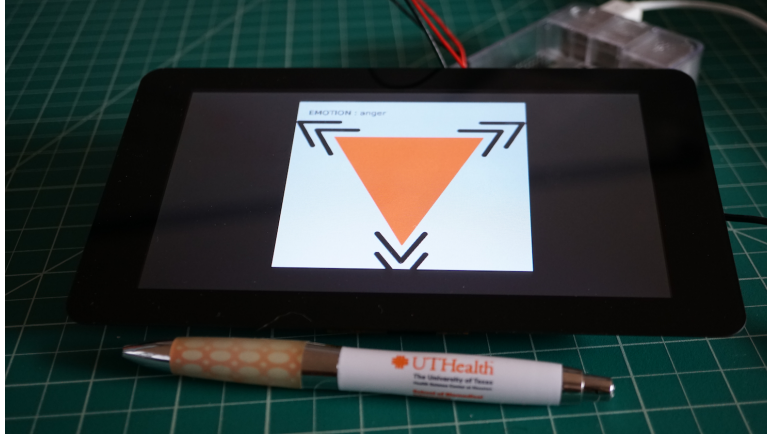
1. Developed the proof-of-concept engine that harnessed VEO to allow for querying and interpretation of emotions using an application programming interface (API)
2. Tested the VEO-Engine's functionality to query and perform reasoning for emotions

To test the software library, we used a Raspberry Pi 3 Model 3 board with Raspian version 9. Specific to Raspberry Pi, the device was also connected to 7" touchscreen display with 800 x 400 pixel screen resolution. The VEO-Engine was deployed to the device, and we executed sample tests through the command line to assess both the visualization query and the emotion reasoning of the library.

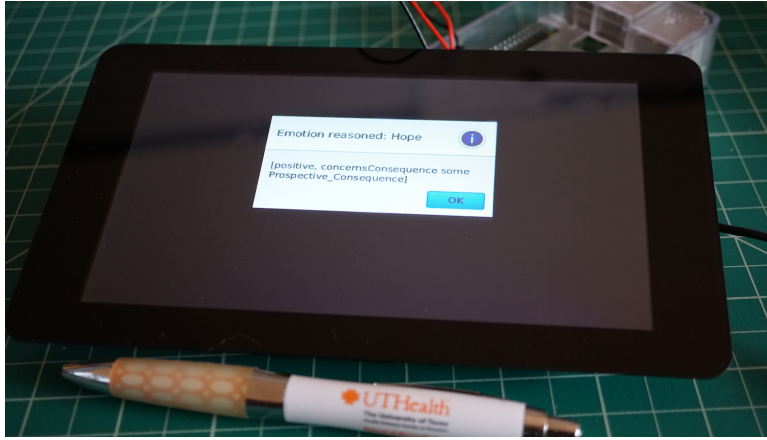
Aside from the input parameters we provided through the command line, the entire library was executed locally on the Raspberry Pi device and performed its functions without any connection to external software services.

Through a command line input for a specific emotion, the VEO-Engine queried for the corresponding image file and displayed a sample window showing that the visualization was linked to the emotion.

Figure 5.12 shows *anger* displayed from the VEO-Engine on a Raspberry Pi device.



**Figure 5.12:** Touchscreen device displaying results of a visualization query for the emotion *anger*.



**Figure 5.13:** VEO-Engine performed a reasoning task based on parameters for the emotion of *hope*.

We tested the VEO-Engine’s reasoner by feeding a string of data describing an emotion. To test, the input required:

$$reason [positive|negative] [concept\_property_n]$$

For example, the input parameters of *reason [positive] [concernsConsequence some Prospective\_Consequence]* was revealed to be the emotion of *hope*. Figure 5.13 displays the result of a sample parameter input for *hope* to demonstrate the reasoning capability of VEO-Engine on a small device.

### 5.3 EXAMINING USABILITY AND IMPACT ON THE USERS' HEALTH BELIEFS

We assessed the usability of the conversational agent for HPV counseling (referred to as “Beverly”) on individual users. One group involved parents with at least one child under 18 and another group involving young adults who were under 25, an age group that the vaccine is still effective for. These two groups are potential users for whom the HPV vaccine counseling agent could target. In the following sections, we describe some early assessment through a simulated agent to gain some insight to develop an automated tool (some of the components discussed earlier were the result of that insight) and also attain some evidence on the probable success of this system if it were to be used in a clinical environment. Section 5.3.1 looks at an early pilot using Wizard of OZ protocol (see Chapter 4) with parents with at least one child under 18. Section 5.3.2 looks at an early pilot with the same protocol but with young adults at a major Texas campus. This specific trial collected more detailed data, like perceived beliefs of the HPV vaccine and data from validated usability surveys.

#### 5.3.1 EARLY USABILITY ASSESSMENT OF CONVERSATIONAL AGENT

Our aim is to gather preliminary assessment knowledge of a vaccine-centric conversational agent (CA) to help refine our idea of utilizing an automated CA for HPV counseling at a clinical environment. For our early assessment and data collection process, we employed the Wizard of OZ protocol that simulates speech interfaces with a potential user who thinks they are interacting with an automated machine or robot (Fraser & Gilbert, 1991). It allows us to evaluate the usability and acceptability of the system rather than to measure the quality of an entire system.

We proposed the following questions:

- How would parents with a child under 18 assess the usability (ease of use, efficiency, and expected capabilities) of a voice user interface for HPV vaccine counseling application?

- Would the parents' vaccine hesitancy have an impact on the usability of an HPV vaccine conversational agent?
- What are the features and requirements that users desire in order to be feasible?

Our study was approved by UTHHealth's Internal Review Board and conducted from February to July of 2018. Flyers were posted across the campus to advertise for participants. During that time period, we recruited 18 participants, who were adults with at least one child under age of 18. This is primarily because the HPV vaccine is an adolescent vaccine that is administered between the ages of 11 through 18, and the parent is the decision maker for the child. Each participant was escorted to a private room by the data collector assistant and completed a pre-assessment survey that included basic information about the subject and a Parent Attitudes about Childhood Vaccines (PACV) survey (Opel et al., 2013) that measures vaccine hesitancy. Adjacent to the room, another researcher, the operator, was seated with the desktop application and the dialogue script.

After completing the pre-assessment survey, the experiment started, and the participants went through the simulated automated counseling system with the operator coordinating the interaction through the guidance of the dialogue script. At the end of the simulated counseling session, we administered a usability survey voice user interfaces provided by (Cohen et al., 2004) that had three questions pertaining to the aforementioned usability variables, and we collected the free text comments from the participants which were later segmented by positive and negative comments. Out of the 18 participants, one refused further participation and one experienced technical difficulty, and overall, the final count of participants was 16. Of these 16 participants, 6 spoke English as a non-primary language. There was an equal number of healthcare professionals or researchers and non-healthcare professionals or researchers. 9 of the participants had a graduate degree. Most people have children below 10 years of age.

The PACV survey measured parent attitudes about childhood vaccines. Each answer had a value and a raw score was obtained for all of the questions. That score was then converted to a score that

measured vaccine hesitancy. A score of 0-50 represents not vaccine hesitant, 50-80 represents vaccine hesitant, and 80-100 represents very vaccine hesitant. All 16 of the participants were considered not vaccine hesitant, which prohibited us to measure the impact of vaccine hesitancy on the usability of the voice interface. This lack of vaccine hesitancy is likely the result of recruiting at a health science center, where individuals are presumably better informed on vaccines.

On a scale from 1-7, there was an average score of 5.4 ( $\sigma=1.59$ ) for ease of use and an average of 4.5 ( $\sigma=1.46$ ) for the expected capabilities. Most of the participants relatively disagreed with the statement that the system was quick and efficient with an average score of 3.3 ( $\sigma=1.85$ ) on a scale from 1-7\*. On average, participants relatively agreed that the system was easy to use and had the capabilities that they expected. However, the perceived slowness was attributed to how fast the remote operator responded to the user's utterance, and in an automated setting would be less of an issue.

The Pearson's correlation coefficient ( $r$ ) was used to quantify the linear correlation between two variables in order to investigate if the score of one variable affected the other in a linear fashion.

Correlation coefficient is a continuous number ranging from -1 to 1, with -1 stands for perfectly negative linear relation, and 1 stands for perfectly positive linear relation. For interpretation purpose, we classify the absolute values of correlation coefficients of less than 0.30 as "small or no correlation", values of [0.30, 0.50] as "weak correlation", values of [0.50, 0.70] as "moderate correlation", and values of [0.70, 1.00] as "strong correlation". Also, we calculated the Fisher's z transformation to obtain a normal distribution for our correlation calculation.

As a result of this correlation analyses, we found the weak correlation between the ease of use and the efficiency of the system with an estimate of 0.34 (95% CI [-0.18, 0.72], with two-sided p-value of 0.197), the moderate correlation between the ease of use and the expected capabilities of the system with an estimate of 0.63 (95% CI [0.08, 0.82], p-value=0.024), and the moderate correlation

---

\*7-strongly agree, 6-agree, 5=agree somewhat, 4=neither agree nor disagree, 3=disagree somewhat, 2=disagree, 1=strongly disagree

between the efficiency and the expected capabilities with an estimate of 0.55 (95% CI [0.20, 0.86],  $p\text{-value}=0.007$ ), where the 95% CI were computed by proper use of Fisher's z transformation (Fisher, 1915). In summary, these analyses suggested that the score given for the expected capabilities has a moderate effect on the scores given for the ease of use and efficiency.

The written comments from the participants can be separated into two broad categories, negative and positive. The negative comments mainly concerned five things: the response time, the repetitiveness, the lack of visuals, the need to humanize the system, and the inability to answer all questions. The biggest concerns were regarding the response time and the repetitiveness. Most people wrote that the system needed to improve its response time and that it repeated the same points. People also stated that adding graphics would help with the overall look, improve interaction, and make it less uncomfortable. Some people also stated that they would prefer speaking to a person or a system that sounded more human-like. The positive comments mainly regarded four things: the interactivity, how informative it is, the accessibility, and the clarity of the system. Most people enjoyed the interactive aspect of the system and how it understands them and encourages them to ask questions. Many people also stated that the system provided useful information on the HPV vaccine and reinforced the important points. Not only did the system offer useful information, it also stated the information clearly, had a clear voice, and was very straightforward and easy to operate.

### 5.3.2 EXTENDED USABILITY ASSESSMENT OF THE CONVERSATIONAL AGENT

Again, we utilized the Wizard of Oz (Fraser & Gilbert, 1991) where the dialogue exchange is handled by a drone operator of the conversational agent (designated as "Beverly"). Potential participants were young adults above the age of 18 to 26. The participant was offered an explanation of the study (design, procedures, and risks) and was given time to ask questions. Each participant was told that their participation is voluntary, and they can withdraw from the study at any time without negative consequence. If they agree to participate, the study personnel will obtain informed consent from

them. The participant entered a designated observation area on TAMU campus with a conversational agent device (simulated – Wizard of OZ). The participant was left with the conversational agent and the conversational agent initiated conversation with the user. For 15-30 minutes, the patient and conversational agent engaged in discourse and discussion about the HPV vaccine. The drone operator utilized a script designed by the study personnel to counsel on the HPV vaccine (Amith et al., 2019b) (See Chapter 4). After the session of the simulated counseling information session, the study staff administered a few surveys for the participant to complete.

The surveys were provided in English and consisted of three instruments – System Usability Survey (SUS) (Brooke et al., 1996; Brooke, 2013), The Speech User Interface Service Quality (SUISQ) (Polkosky, 2008, 2005), and the Carolina HPV Vaccine Attitude and Belief Scale (CHIAS) (McRee et al., 2010a). SUS is a validated industry standard to provide a “quick and dirty” scoring for usability, providing a rating between 0-100 derived from 10 survey items (Appendix E). The SUS scoring can be interpreted by associating a school grade (i.e. 80-89 = B, 70-79 = C, etc.) or associating with an adjective rating (Sauro & Lewis, 2016; Bangor, 2009).

SUISQ is a 25 question survey developed specifically for interactive voice response applications based on four factors: User Goal Orientation (8 items), Customer Service Behaviors (8 items), Speech Characteristics (5 items), and Verbosity (4 items). Each item is on a Likert scale between 1 to 7. The factor scores are the mean of their items and an overall score is the mean of the factor scores (Verbosity mean has to be reversed – subtract from 8). User Goal Orientation describes the “system’s efficiency, user trust, confidence in the system, and clarity of the speech interface”. Customer Service Behaviors involve “friendliness and politeness of the system, its speaking pace, and its use of familiar terms”. Speech Characteristics refer “naturalness and enthusiasm of the system voice”, and Verbosity is the “talkativeness and repetitiveness of the system.” (Lewis & Hardzinski, 2015).

CHIAS is a 16 item survey relating to the four factors of the health belief model (Harms, Effectiveness,

Barriers, Uncertainty) for the HPV vaccine. This survey has demonstrated stability of the factors to describe HPV vaccination attitudes over time. With the exception of the Effectiveness factor, the other three factors have been attributed to parent HPV vaccination intention and predicted HPV vaccination utilization of parents (Brewer et al., 2011). The CHAIS is adaptable for young adults and have shown to have similar validity to the parent version (Nicol et al., 2016; Dempsey et al., 2014). Two recent studies utilized this variation of CHIAS on samples of a young population under 25. Kamimura et al. (2018) utilized the Risk and Denial, and Benefit factors of the survey to measure differences between young Vietnamese and United States college students. For Hanson et al. (2019), they focused and measured on the Effectiveness, Uncertainty, and Harms factors of the HBM for their survey study.

One aspect where we were interested in exploring was how this agent would fare with similar systems that employed a voice interface. This would give us an idea of how receptive and how effective a conversational agent for HPV vaccine would be to future users. Furthermore, we also wanted to know the impact of this type of software would be on vaccination uptake. The health belief model has some predictability towards cue to action, namely vaccination uptake (Janz & Becker, 1984). If there is some early evidence that a conversational agent may have on their beliefs it maybe encouraging enough to further pursue this area in later studies. Therefore, we propose the following questions:

- What is the overall usability of “Beverly” (i.e. simulated conversational agent) compared to the usability of most interactive voice user interfaces?
- What is the association of the usability of “Beverly” with the users’ perceived beliefs of the HPV vaccine?
- Does “Beverly” improve the users’ perceived beliefs of the HPV vaccine?

We conducted the experiment on individuals of young adults who attend Texas A&M University-College Station. The sample size was 25 who responded to a call for recruitment on the campus,



coordinated by our collaborators from Texas A&M University. The inclusion criteria were open to any individual. Overall, among the sample, 11 were females and 14 were males, and the average age is 20. Ethnic breakdown was 19 white, 3 black, 2 Hispanic, and 1 Asian. 8 had the HPV vaccine, 8 did not have the HPV vaccine, and 9 don't know. 1 user who experienced technical difficulties was removed from analysis ( $n = 24$ ).

We collected their survey responses from the aforementioned instruments. Appendix E list the various variables from the survey. For four questions of the HBM survey, we reversed the scale to match the direction of the HBM variables - higher scores were indicative of less receptiveness to the HPV vaccine. For the SUS survey, we neglected the question "*I think I would like to use this system frequently*". According to Lewis & Sauro (2017) analyses, removal of this question would not have a major impact on the statistical validity of the SUS final score. To compensate for the neglected question, we adjusted the SUS calculation based on what is described by Lewis & Sauro (2017). This was done because we do not foresee this agent being used everyday as a consumer tool. Essentially, we envisioned the agent to be situated in a clinical environment while the patient is waiting for their health care provider. Which means they will most likely encounter the system at least once. All of the statistical analysis was completed through SPSS v25.

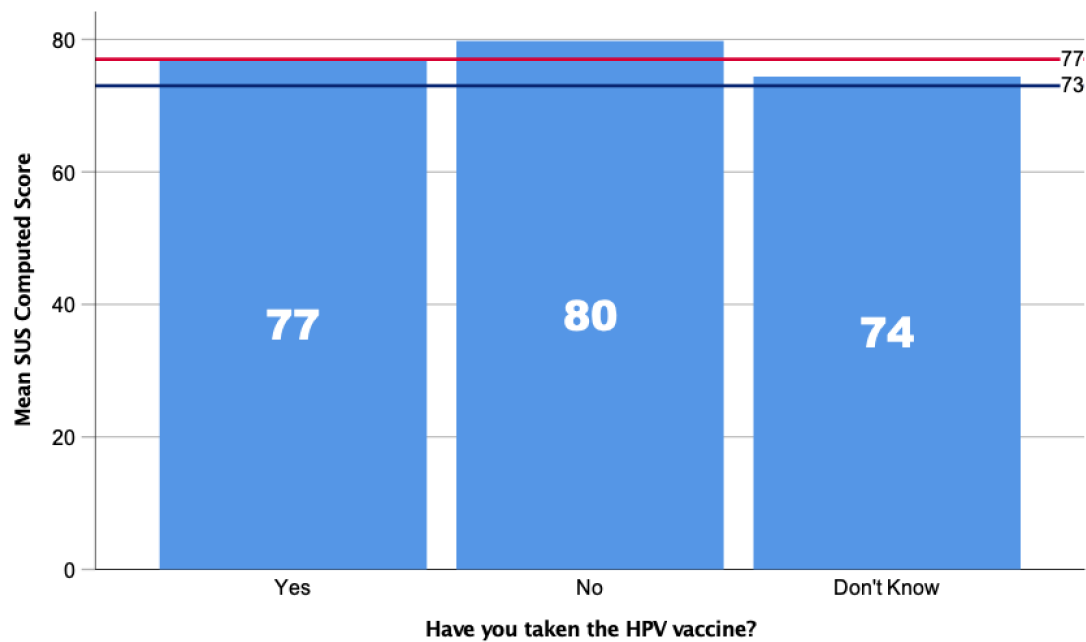
For the primary question for overall usability, we computed one sample t-test against an average SUS score for voice interfaces,  $\mu = 72$  ( $n = 233$ ) (Bangor, 2009). We performed a one sample t-test to compare the mean of the SUS score with the reported industry SUS mean for voice interfaces. The analysis produced a t-statistic of 1.627, with a probability score of 0.059. Overall, there is 94% confidence that the system has a score above the industry mean of 73. This would reveal that the conversational agent had an on-par or a slightly above average SUS score with industry-based interactive voice interface systems. We also looked at the differences between users by vaccination status (See Figure 5.14).

The SUS overall score was higher among those that never had the vaccination ( $\mu = 80$ ) compared

to those that did have the HPV vaccine ( $\mu = 77$ ) and those did not know if they had the HPV vaccination ( $\mu = 74$ ).

We also looked at the attributes of individuals in relation to the SUS score to ascertain any association with the SUS Score. A Fisher Exact Test ( $p = 0.54, 12.42$ ) indicated a statistically marginally significant relationship between vaccination status and SUS score. None of the other attributes such as school classification, parents' income, etc. indicated any significance with the SUS Score.

While the SUS survey has a history of being a valid measurement tool for quick usability scoring, the SUIQ survey has yet to be as standardized as the SUS. We administered the SUIQ survey in conjunction with SUS, as recommended by [Lewis \(2016\)](#). We looked at the correlation between SUS and SUIQ to ascertain any significant relationship. For the most part, there appear to be positive correlation between SUS Score and SUIQ score, moderately strong, and statically significant



**Figure 5.14:** SUS Scores among different vaccination statuses. Red line is the average of all the participants. Black line is the baseline mean measure reported for interactive voice interface systems reported by [Bangor \(2009\)](#).

( $r=0.486$ ,  $p=0.016$ ). With the User Goal Orientation and Verbosity factors of SUI SQ, there also appear to be a strong, positive relationship with the SUS score ( $r=0.508$ ,  $p=0.005$ ). The same is said for Verbosity with SUS Score ( $r=0.627$ ,  $p=0.000$ ). There is also moderately positive correlation with Customer Service Behavior Score with the SUS score ( $r=0.438$ ,  $p=0.014$ ).

For the next group of analysis, we investigated any effect of usability on the participants' beliefs of the HPV vaccine. Looking at those who reported to not had the HPV vaccine and did have the HPV vaccine, we calculated the Spearman rank correlation. Among those that reported to not had the HPV vaccine, there was a strong inverse relationship (i.e. lower CHAIS rating indicates a greater propensity for the HPV vaccine) between SUI SQ and the effectiveness construct, which was statistically significant ( $r_s = -0.711$ ,  $p=0.037$ ). There was also a strong inverse relationship between Verbosity factor of SUI SQ and the uncertainty construct for the CHAIS health belief model, which was statistically significant ( $r_s = -0.874$ ,  $p=0.05$ ). Among those that reported yes to the HPV vaccine, there was a strong inverse relationship between SUI SQ and the barrier construct, which was statistically significant ( $r_s = -0.639$ ,  $p=0.044$ ). Also a strong inverse relationship between SUI SQ and the effectiveness construct, which was statistically significant ( $r_s = -0.655$ ,  $p=0.039$ ). In addition there was a strong inverse relationship between Verbosity factor of SUI SQ and the effectiveness construct, and statistically significant ( $r_s = -0.764$ ,  $p=0.014$ ).

Lastly, we examined if there is any possible improvement in the users' perception of the HPV vaccine compared to samples from published studies conducted by Kamimura et al. (2018) and Hanson et al. (2019). We re-scaled our collected CHIAS HBM data to match the scales for Kamimura's and Hanson's CHIAS data – five point and 11 point scale respectively. Then we computed a one-sample t-test with the means of the CHIAS factors with the CHIAS factors of the aforementioned published studies. Table 5.8 summarizes the our results where the green signifying statistical significance, yellow with marginal statistical significance and red with no statistical significance.

There appear to be better beliefs and perceptions with Risk and Denial ( $\mu = 1.41$ ,  $p=0.00$ ), Perceived

**Table 5.8:** Comparing means of CHAIS constructs from the intervention and published studies. Lower values for CHIAS indicate positive attitudes in favor of the HPV vaccine.

CHIAS HBM Construct	Agent+CHIAS	CHIAS	p-value
Perceived Barriers	1.94	2.33 (n=437)(Kamimura et al., 2018)	0.09
Risk and Denial	1.41	2.00 (n=437)(Kamimura et al., 2018)	0.00
Perceived Harms	1.92	3.5 (n=108)(Hanson et al., 2019)	0.00
Perceived Effectiveness	3.82	4.3 (n=108)(Hanson et al., 2019)	0.26
Perceived Uncertainty	2.13	4.8 (n=108)(Hanson et al., 2019)	0.00

Harms ( $\mu = 1.92$ ,  $p = 0.00$ ) and Uncertainty ( $\mu = 2.13$ ,  $p = 0.00$ ) among the individuals who experienced “Beverly”. There was some improvement if accounting for the marginally statistical significance with Perceived Barriers ( $\mu = 1.94$ ,  $p = 0.09$ ). There was no evidence of improvement with Perceived Effectiveness despite better rating ( $\mu = 3.82$ ,  $p = 0.26$ ).

Overall, based on the initial analysis, we could conclude that the possibility of using a conversational agent for the HPV vaccine could be as good the industry standard for usability of a voice user interface. And that the usability could have some correlation on the health beliefs of its users as it relates to the HPV vaccine. While these results are preliminary at best, and lack a larger sample size, there is some promise that deploying a tool for a health care environment could be feasible. However, additional research is needed to further validate our assumptions.

#### 5.4 SUMMARY

This chapter details the assessment of the various components for a proposed ontology-driven architecture for HPV vaccine counseling. We reviewed the intrinsic and extrinsic quality of ontologies for VISO-HPV, VAXMO, PHIDO, and VEO using semiotic theory-based measures. The results indicated satisfactory quality when compared to sets of ontologies that are were of the same domain or from a sample benchmark of NCBO BioPortal ontologies. We also evaluated the system functionality of ontology-driven software components that used some of the aforementioned ontologies as their

core component. The ontology-driven software engines (VEO-Engine, COO, and FOQUS) were able to perform with functional tests to adhere to tasks for which they are designed for the agent. Lastly, we also gathered user data on the perception of a conversational agent and possible effect on their beliefs of the HPV vaccine. The results show some influence of the usability of the voice interface with specific health belief model constructs for the HPV vaccine (e.g. perceived effectiveness and uncertainty) among those that have reported to have had the HPV vaccine and those that reported to never having the HPV vaccine. There is some evidence that a conversational agent can improve the perception and beliefs of the HPV vaccine when we compare to existing studies that measured health beliefs of the HPV vaccine among the young. Overall the domain, application, and interaction view of the system architecture provided the possible realization of the development and deployment of an automated conversational agent in a live environment. In the next and final chapter, I will discuss the implication of our work and limitations (i.e. our future direction).

*Perfection (in design) is achieved not when there is  
nothing more to add, but rather when there is nothing  
more to take away*

Antoine de Saint-Exupery

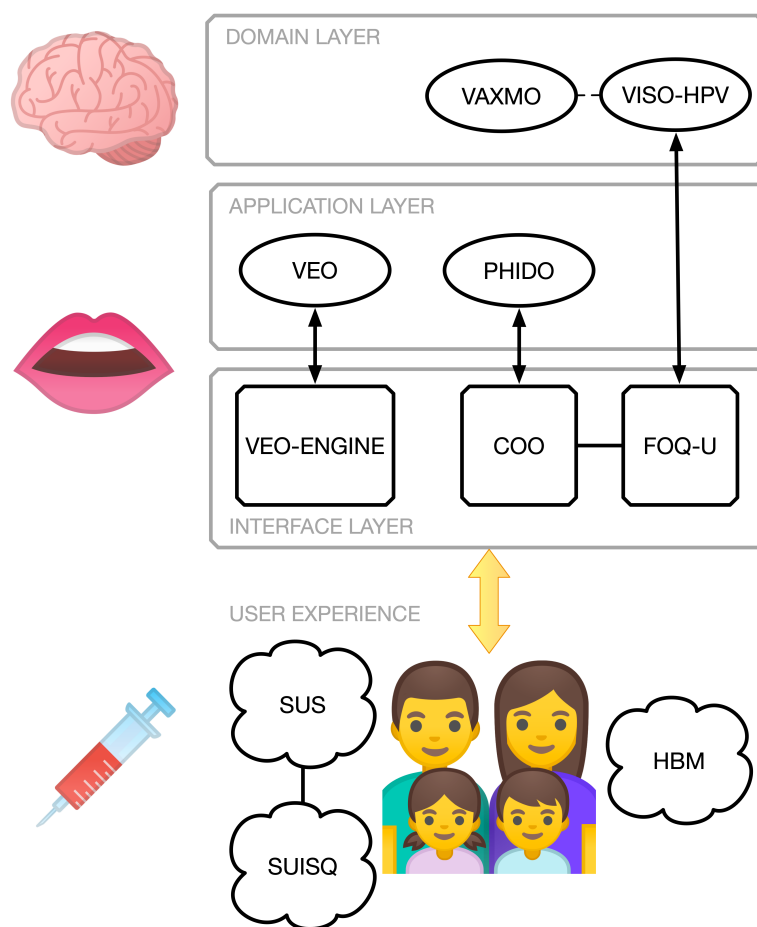
# 6

## Conclusion

We presented an ontology-driven architecture for a speech-enabled conversational agent for the HPV Vaccine. This architecture is modeled using a vertical layered structure, where each layer is facilitated by an ontology artifact that provides the intelligence to interact with the user. Each component was scientifically evaluated. For example, ontologies were assessed using a metric suite against other ontologies, and if they had any unique provision, we assessed the ontology with an appropriate evaluation. Application software components were tested for their primary functionality, ranging from the ability to retrieve visualizations based on an emotion, interpret emotion, answer HPV vaccine-related questions, and manage the dialogue. Aside from the system architecture, we also presented some evidence to support the notion that high usability of an automated spoken dialogue could impact the beliefs of the user's perception of the HPV vaccine. The combination

of the architecture and early results from usability tests provides some feasible support for the development and deployment of an ontology-based spoken dialogue system for the HPV Vaccine. In this last chapter, we discuss some of the results and limitations of this study, and offer some future steps and directions to expand this work.

## 6.1 DISCUSSION



**Figure 6.1:** Layered architecture of the conversational agent for HPV vaccine.

For simplicity, we will refer to the figure from Chapter 1 that outlines the scope and system architecture

of this study (Figure 6.1). We will start by discussing the user experience and user impact involving the usability of the HPV vaccine conversational agent, and the effect it had on their perception of the HPV vaccine (User Experience, Section 6.1.1). Then we will segue in explaining the work related to the ontology-driven interface layer (Interface Layer, Section 6.1.2), and the ontologies related to the application layer (Application Layer, Section 6.1.3). Lastly, we will briefly review the work in the domain layer (Domain Layer, Section 6.1.4). This chapter will end with some notes on contributions and limitations that could inform the future direction of the research in this area.

## 6.1.1 USER EXPERIENCE

### 6.1.1.1 EARLY ASSESSMENT WITH PARENTS

While we received some encouraging positive comments, there were some useful suggestions that could lead to an improved user experience. The responsiveness issue was an important aspect. Because we were copying and pasting or typing responses to be transmitted wirelessly that may have had a slight latency impact on perceived responsiveness. It also highlights that if the system were to be automated and utilize artificial intelligence components, the system would need to be relatively quick in responding to the user's utterance. Another issue was better usage of graphics or visualizations to complement the dialogue. Lastly, the aim of the speech system is to alleviate some of the communication challenges at a clinical environment. Some of the users expressed a desire to speak to a human.

Our belief is that the important discussion points, like personal contextual health information as it pertains to vaccines should be handled by the provider to avoid confusion. Throughout the dialogue, we emphasized that user should confer with the doctor for more nuanced and specific information, especially since we envision this conversational agent to be stationed as a kiosk or tablet in the waiting room.

From our study of 16 participants, who were not vaccine hesitant, we determined that parents found



that a vaccine conversational agent was relatively easy to use and had the expected capabilities. Most participants found the agent to be slow and this is mainly due to responsiveness of the Wizard of OZ (WOZ) remote operator. Nonetheless, the feedback highlighted the importance for automated vaccine conversational agents to be responsive with utterances and with the suggestions from users, therefore, we have collected future features and improvements to better develop an automated CA.

#### 6.1.1.2 EXTENDED ASSESSMENT WITH YOUNG ADULTS

This second attempt at the Wizard of Oz experiment was performed on young adults who are also potential targets for the HPV vaccine, considering the recent FDA approval for its effectiveness through age 45 ([US Food and Drug Administration, 2018](#)). Contrasting from the previous attempt, we administered a more detailed usability survey — System Usability Scale (SUS) and Speech User Interface Service Quality (SUISQ). SUS is a general usability measurement scale based on 10 Likert scaled questions that has proven reliability. While SUIQ is a more robust usability survey for voice user interfaces, it has yet to achieve the reliability of the SUS. Together they may provide further validation support for SUIQ ([Lewis, 2016](#)). We also administered a health belief model survey Carolina HPV Immunization Attitudes and Beliefs Scale (CHIAS) that was tailored for young adults ([Dempsey et al., 2014](#)) to see of any correlation between the health belief constructs and usability variables. We also compared the means for each of the constructs with previous published studies that used this modified survey for a younger population ([Kamimura et al., 2018](#); [Hanson et al., 2019](#)).

The overall usability (i.e. satisfaction, effectiveness, and efficiency), measured through the SUS score, indicated that the conversational agent had average to slightly better than the mean of 72 among 223 interactive voice interface system, with an average SUS score of 77 (94% confidence). To add, there was marginally statistical significance relationship with the SUS score and the HPV vaccination status. There were slight variations of the SUS score among the different users, with

those that had the HPV vaccine (with 77) and those that do not know if they had the vaccine at 74. The individuals that never had the HPV vaccine rated the system high at 80. The score from this group is meaningful as this tool could be effective for individual users who never had the HPV vaccine and could perhaps inform and encourage them better than other paper-based methods. With a small sample, we can conclude that there is evidence for those never had the HPV vaccine that our conversational agent (while simulated) has strong usability with the tool. We need to further analyze why those that don't know they had the HPV vaccine scored lower at 74 and perhaps look into ways to improve among this groups of individuals.

Lower values for the health belief ratings indicated a propensity for the HPV vaccine. Therefore, we sought any statistical significance showing an inverse relationship between the CHIAS health belief model (HBM) constructs and usability factors. No significant relationship exists between SUS and CHIAS HBM constructs, but there are some statistical significant relationships between the SUISQ factors and CHIAS HBM constructs. Among the individuals who never had the HPV vaccine, the SUISQ Overall Score has a strong inverse relationship with perceived effectiveness of the HPV vaccine. This would indicate that the participants who never had the HPV vaccine and the liked agent, are more likely to consider the HPV vaccine and believe in its effectiveness. With that same group, the Verbosity factor — measuring how repetitive and talkative the agent is — had a significant inverse relationship with the perceived uncertainty. For background, a lower Verbosity indicated low perception of repetitiveness and talkativeness. This score was reversed for analysis where a higher Verbosity score indicated low perception of repetitiveness and talkativeness of the agent. Perceived uncertainty represents how susceptible they feel on attaining a health issue. In this case if they perceive they are susceptible to the HPV they are likely to engage behaviors to prevent it. What we gather from our results for the relationship between Verbosity and Uncertainty is that the information that was spoken by the system was not deemed insignificant to the user and that the content related to the virus had an effect on their knowledge of HPV.

Of interest, the individuals who reported yes to having the HPV vaccine also shared the same relationship pairing — Verbosity  $\approx$  Uncertainty, and Overall SUISQ  $\approx$  Effectiveness. Also, SUISQ had a significant relationship with Uncertainty and Barrier. The Barrier rating reveals any external challenges that could impede attaining the HPV vaccine. Since individuals reported yes to having the HPV vaccine it is clear they do not have any barriers preventing them since they are vaccinated. The various relationship with several variables are more likely due to them already having the HPV vaccine, since SUISQ overall score was associated with three of the five CHIAS constructs.

When we analyzed the Carolina HPV Immunization Attitudes and Beliefs Scale (CHIAS) user data with CHIAS mean results from other published studies that used the same survey ([Kamimura et al., 2018](#); [Hanson et al., 2019](#)), there is a statistical difference (i.e. improvement) with various means of the constructs for the Health Belief Model. Aside from the perceived effectiveness construct, there was also a marginally statistical difference with perceived barriers construct. In absence of a true control group, there is some evidence of impact on health beliefs and perceptions of the HPV vaccine as a result of the using a conversational agent for HPV vaccine.

The young demographic may have different information needs compared to adults. For example, emphasizing health sexual practices to prevent HPV transmission. Among those that reported not knowing whether they had the HPV vaccine, no significant inverse relationship with any of the variables existed. We need to investigate why this was the case. It might be possible that they have low health literacy, but we cannot be certain unless we measure their health literacy skills. It may also be due to their lack of interest in their personal health, in which case a tailored dialogue intervention may be needed for this group.

## 6.1.2 INTERFACE LAYER

### 6.1.2.1 COO AND FOQUS

The Conversational Ontology Operator (COO) was supported through the use of PHIDO. Using PHIDO as the facilitator for the dialogue engine we were able to demonstrate the use of an ontology to control the flow of the dialogue and maintain the dialogue context at the same time. Three use cases were introduced – communicating one statement related to health information, facilitating the interaction for question answering, and transitioning to the next topic. In all of the use case tests, the engine was able to support the various dialogue interactions.

FOQUS provided question answering abilities to answer sample questions from the simulation logs. It utilizes two variants (string matching and vector-based comparisons) to find matches of salient concepts of the question with the triples of the ontology. Irrespective of the configuration for FOQUS, the question answering system did perform sufficiently in answering the questions from the chat logs collected from our Wizard of OZ experiment, with an accuracy ranging from 0.50 to 0.72 (depending on the variant or the inclusion of partially answered responses). With some promising initial results and a system foundation to build upon, refinement is needed to further improve FOQUS. We may explore natural language generation methods to better improve the transformation of triples to clear and natural answers. However, one limitation of this study is that we may need to factor in the impact of answers being uttered by a machine. For this, we need to assess FOQUS in a live environment with users and test its portability with other consumer ontology knowledge bases. Further discussion of the future directions and limitations of COO and FOQUS is provided in 6.3.

#### 6.1.2.2 VEO-ENGINE

While our results show promise for semantic driven technologies, there are still opportunities for improvement. One would be to allow for synonymous emotion input in visualization queries, for example, *fondness* in place of *love*. To permit this, we would need to expand the ontology to link similar terms with each emotion and then modify the SPARQL queries. These improvements are possible because ontologies are graph-like, and therefore they can be changed easier than, say, a relational database (Allemang & Hendler, 2011).

To perform reasoning functions, the VEO-Engine required structured data input, so for this technology to be further applicable, it must map or translate the noisy contextual information from the human user into structured data. Therefore, if we looked at unstructured, free text from a person's utterances, we would need to parse out the information and then map that information to the appropriate parameters for emotional valences and concept properties to then input into the VEO-Engine. In this scenario, natural language processing might offer a direction.

Our work exemplifies how semantic-encoded emotions could be utilized by software and small devices to assist machines in understanding human emotions. Based on our previous VEO work (Lin et al., 2018), we developed VEO-Engine, a software library that interfaces with the emotion ontology. The VEO-Engine was able to query for visualizations associated with an emotion, and it was able to deduce an emotion based on sample input parameters. The combination of having emotions semantically defined and a software wrapper to interface with the ontology makes semantic web technologies a feasible option for affective computing. In the future, we will look to incorporating this work into conversational agents for health care applications. Specifically, this could enhance how machines react and respond to patients' or health consumers' utterances to improve their outlook and well-being.

### 6.1.3 APPLICATION LAYER

#### 6.1.3.1 PHIDO

The representation of the patient-centric counseling was inspired from the dialogue script and the Wizard of OZ implementation of the script with live participants, which included the interaction logs from their participation. While the application ontology is rooted in real world activity, there will likely be exceptional utterances that we may not anticipate by future user interactions. Currently, our Wizard of OZ experiment is ongoing and future interaction logs may inspire modifications for the ontology.

We stressed that PHIDO is an application ontology, so it may not universally cover the domain of dialogue interaction, nor are there immediate plans to align it with an upper ontology. Currently our focus is on vaccine counseling, but we foresee the possibility that PHIDO could cover patient-centric communication of health information for a variety of topics while being grounded in some behavioral theory. Also of similar importance, as indicated by the domain coverage, PHIDO would need to model additional Speech Tasks since it only represented 6 types, which is not extensive.

Dialogue management is bifurcated into dialogue flow and dialogue context components (Jokinen & McTear, 2009). Most of what we discussed for PHIDO facilitates dialogue flow with some minimal contextual information (e.g. the Utterance class's *hasBeenSaid*). Ideally, we hope in the future that PHIDO will encompass management for dialogue flow and contextual information for an ontology-driven approach.

We derived an application ontology for dialogue management called Patient Health Information Dialogue Ontology (PHIDO) that is based on our on-going Wizard of OZ experiments conducted at University of Texas Health Science Center at Houston. This application ontology is intended to be used in a prospective dialogue engine for embedded and mobile devices that will automate a counseling session for the HPV vaccine, a vaccine that has dramatically low coverage among the

population. Our initial qualitative results based on the semiotic metric suite indicated that PHIDO is of comparable quality to NCBO Bioportal ontologies. Our current activity is to develop the software engine that will harness PHIDO to be deployed in machines, and to link a lightweight ontology-based question and answering system to the dialogue manager. We foresee that our work will demonstrate and contribute to the usefulness of semantic web and ontology technology to power patient-centric conversation for health information.

#### 6.1.3.2 VEO

In the future, we could expand VEO by creating nuances within certain emotion types—for instance, fear-like states can range from those that are mild (e.g., concern) to those that are intense (e.g., terror). These types of states could be included as subclasses in the ontology. We also intend to expand the terminological space with some of the affective terms found in WN-AFFECT. Additionally, we could add instances in the future that represent an individual user's emotions.

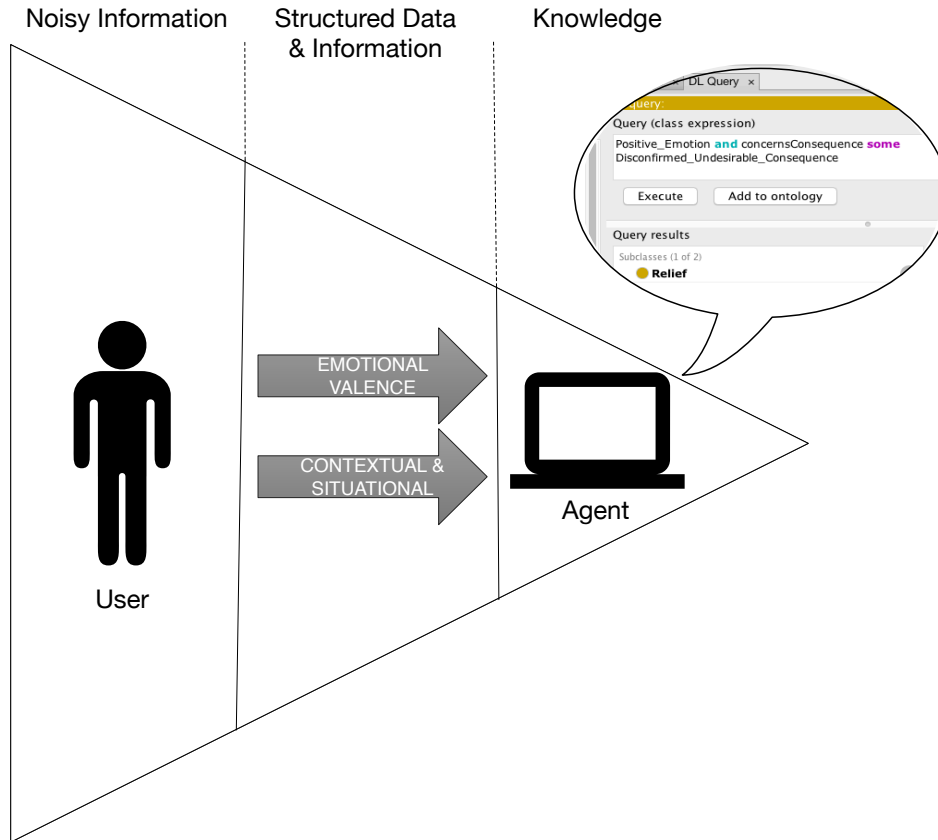
Overall, the survey results validated the accuracy of our emotion visualizations. More people agreed than disagreed that the image matched the emotion displayed for 17 emotions (with 16 out of these 17 emotions found to be statistically significant), and vice versa for eight emotions. However, only for the three emotions of distress, reproach, and resentment did people prefer the incorrect emotion-image pair to the correct one. One reason the incorrect emotion-image pair was preferred for distress could be due to its name—distress and sadness have slightly different connotations, and if we had used the name “sadness”, perhaps the percentage of people that agreed with our visualization would be higher. After all, even though people thought that the image for fear represented distress (in the incorrect emotion-image pair), they still confirmed that the image for fear was accurate at a high rate (65.0%).

Additionally, in future studies and from the findings of the survey results, it would be helpful to further investigate the eight emotions that did not support our visualizations by comparing them

to different incorrect emotion-image pairs. This could allow us to understand whether the specific randomly-chosen incorrect pair had any influence on our results or if they still hold with different pairs used. If so, these results can inform us in regard to editing our visualizations so that they are more representative of each emotion. Our research also does not consider the use of motion, which could enhance the visualizations in the future.

In the future, this study could permit machines to utilize the VEO to interpret and understand emotions, with the purpose of improving interaction with human users, such as patients. For clarification, recall that ontologies are artifacts of encoded knowledge to help machines understand domain concepts and the relationships between them. Codifying affective knowledge would help intelligent agents, specifically conversational agents, to understand the underlying emotions during their interactions with humans. Looking at an emotion like love, which according to the OCC model, contains positive emotional valence involving the appraisal of some aspect of an object, or anger, which contains negative emotional valence relating to someone's actions and the subsequent outcomes of the actions. A software agent can potentially capture contextual information and emotional valence data, and through the use of descriptive logic queries, reason what the user is feeling or expressing (see Figure 6.2). The use of ontologies to define emotions for machines and then comprehend the emotions of users makes this possible. Further research could investigate processing of the user's emotions from utterances or other modalities of expression. This would also include developing the software that interfaces with the ontology and employing it in conversational agents.





**Figure 6.2:** Interpreting the user's emotional information.

#### 6.1.4 DOMAIN LAYER

##### 6.1.4.1 VACCINE INFORMATION STATEMENT ONTOLOGY

Based on the analysis of the Burton-Jones scores, our initial HPV ontology demonstrating promise. The syntactic score (0.76) demonstrates that the high machine readability of the ontology due to correct use of syntax (lawfulness), and lack of total utilization of OWL features (richness). However, for this specific ontology and its use-case of only representing vaccine patient knowledge, some of the features may not be necessary, but as we continually develop VISO for HPV, there may be knowledge that may require other OWL features to better model the vaccine information. The

semantic quality is also relatively high (0.93) resulting from low use of repetitive terms (consistency) and low use of ambiguous terms (clarity). Overall, the scores are quite promising for initial work, but evaluating the accuracy and expanding the use of OWL-based features are needed to fully evaluate the quality.

Creating the ontology posed many challenges. Even though there is already an existing VISO representation from which the HPV ontology is based, a few concept classes and subclasses had to be created or modified in order to better represent some knowledge triples. For example, a subclass called “Adjuvant” was conceptualized to describe specific ingredients added to vaccines, like monophosphoryl lipid A which is added to Cervarix, one of the licensed HPV vaccines. Another example, the super class “DNA Virus” to describe Human Papillomavirus (HPV) was also created to elaborate on the type of virus that the HPV vaccines target. Another difficulty that we experienced was determining whether to include some of the information found pertaining to HPV in the ontology such as side effects, contraindications, and mechanism of action. Ultimately, any information determined to be beneficial and helpful to the patient was included. Additionally, some of the medical terms in the VIS documents are subject to interpretation and could be ambiguous to patients, such as the terms “mildly ill” versus “moderately ill”. Most patients will be unable to discern the difference between the degrees of illness; therefore, it can become misleading to patients.

As for the accuracy quality, we will need to obtain feedback from subject matter experts to improve the overall quality. We are investigating the possibility of growing the knowledge base with any information that is lacking; for example, inclusion of additional patient-level education about the cancers caused by HPV, such as cervical and oropharyngeal cancer. While the initial knowledge base included basic information regarding HPV-related cancers, additional information is needed to improve the robustness of the knowledge base. Encoding extended information like these into the knowledge base can help raise awareness for the need of HPV vaccination from a cancer-awareness perspective. A neglected aspect, yet of great interest to us, is to go beyond encoding facts and information

about HPV vaccine and consider incorporating multimedia content to enrich the knowledge base. Ontological knowledge bases could indeed link the various concepts with complementary multimedia content, like video and images, to deliver dynamic information through the patient-centered interactive agents. In addition, there is also evidence that storytelling may have an impact to educate and raise awareness of vaccination (Cunningham & Boom, 2013). Whether it would be specific stories linked to certain vaccine concepts or a method to semi-automatically generate a story, the implementation of storytelling components in a machine-readable knowledge base is of interest and will be further explored.

#### 6.1.4.2 VACCINE MISINFORMATION ONTOLOGY

The Vaccine Misinformation Ontology (VAXMO)’s purpose is to catalogue and analyze vaccine misinformation that has been one of the drivers for low rates of vaccination rates worldwide. Ontologies benefit from reusing other ontologies. We have utilized an existing model of misinformation, Misinformation Ontology (MO), to address anti-vaccination information. In addition, we have utilized an innovative approach using nanopublication (which is generally used for scientific assertions) for linking common false assertions or theories about vaccines (i.e. “vaccines causes autism”, “government created weaponized Ebola vaccines”, etc.). Yet, this poses some difficulty - lack of Protégé support and manually editing the ontology artifact. This may inspire us to investigate the possibility of developing a Protégé plugin that provides an interface to view and edit the nanopublication segment of VAXMO. With some modifications, we constructed the ontology based off of the Misinformation Ontology and extended some of its concepts from an existing survey literature. While MO is specifically designed to model false intention and not misfacts, as stated by the original authors, we further extended the ontology to utilize nanopublication graph structure to store and represent false assertions about vaccines. The current representation of VAXMO is encoded in OWL with only the class-level fleshed out and with some conceptual gaps.

Noted earlier, there have been various studies that focused on content analysis of misinformation and myths of vaccines in the public health domain. Some of the literature can help furnish additional concepts to further expand VAXMO, which could help model and understand the features within anti-vaccination information domain.

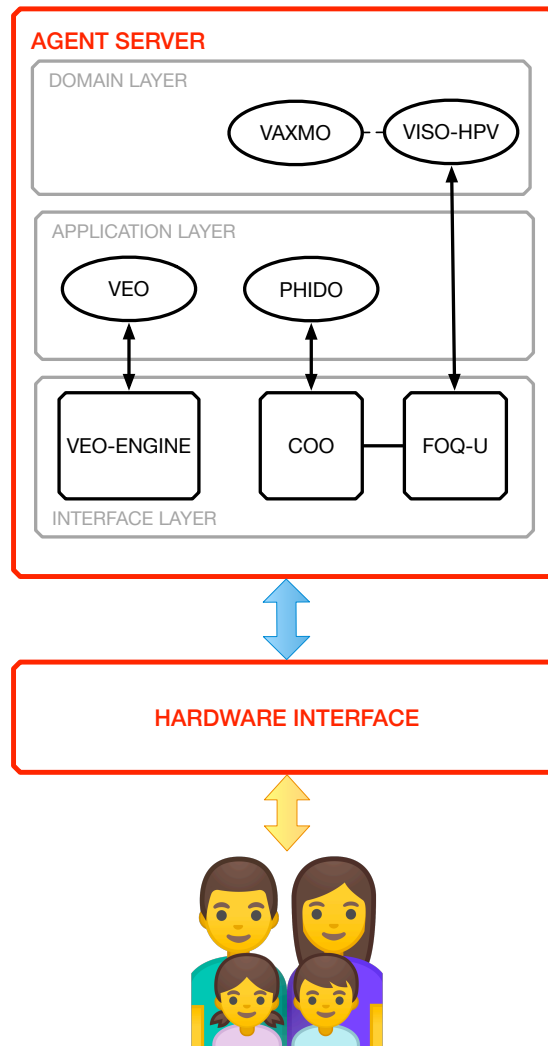
While VAXMO is of better quality than NCBO Bioportal ontologies, there is still some more work needed to expand its conceptual domain space for anti-vaccine information. Also, we have described a future use-case that aims to detect misinformation about vaccines, and we plan on reporting on our findings in a future study.

We assume that the impact of this work could lead to applicable uses of semantic web ontologies for public health informatics and future informatics tools that can assist researchers to understand and address health misinformation in the post-modern era.

## 6.2 CONTRIBUTION

Health informatics is defined as a “the interdisciplinary study of the design, development, adoption, and application of IT-based innovations in healthcare services delivery, management, and planning” (Procter, 2009). The ontology-based architecture described in this dissertation introduces a design for conversational agents for HPV vaccine where each major component - knowledge domain layer, the application layer and the interaction layer - all have a core ontology that facilitate the operations of agent. This would open a whole new avenue for IT integration into the healthcare space and thus expand the field of health informatics.

Without either the PHIDO, VISO-HPV, VAXMO, and VEO, the agent will not be able to direct and track HPV counseling with the user, the question-answering system will not be able to provide triples to answer and form a response, and the visual interface would not be able to provide the utterances an emotive facade or a means to determine and reason the emotion of the user. Each of



**Figure 6.3:** Hardware deployment for a conversational agent. Red highlights the hardware components.

the major components as described in early chapters have shown to preform their desired function. Also, because the software components are domain independent, this architecture was not designed to tightly coupled with the vaccine domain. The domain is defined in the ontologies itself and be all it would require to port to a new domain is to replace the ontology with another. Overall this architecture is a reusable framework for health-based conversational agents.

Aside from the architecture, we were also able to substantiate the possibility of affecting the user's health belief for the HPV vaccine using a simulated conversational agent (the end integrated product). The initial assumption was that patient and provider counseling on vaccine information is able to improve uptake dramatically, but with the burdens placed on providers this is not optimal across every clinic. The results of the study are not only encouraging for vaccination uptake with conversational agents, but it may open new venues to apply this same approach in other important spoken health opportunities, like genetic counseling or PrEP education.

Another important output of our work was the development of a patient-level vaccine knowledge base that can be utilized in consumer-facing software tools. Both VISO-HPV and VAXMO have both been integrated in a components-related conversational agent or demonstrated in use cases where it may be used for analytical purposes (Amith & Tao, 2018). We hope that our work could open the possibility of more patient-centric knowledge bases that can be shared and reused and power the intelligence behind software agents.

Another outcome of our work is the formalization of the dialogue interaction for HPV vaccine counseling. Previous researchers have outlined some high-level architecture and heuristics for discussing vaccines for patients (Opel et al., 2013, 2012, 2015). The initial script provided a framework but also a specific dialogue and was based on constructs of the health belief model. This script was tested in two experiments which showed to be relatively successful with users. The dialogue later served as a template to create scaffolds to within PHIDO, and within PHIDO we can link the utterances to knowledge bases and categories related to the health belief model. In addition, PHIDO offers the opportunity to tailor the dialogue when we begin implementing rules using the Semantic Web Rule Language (Horrocks et al., 2004; O'Connor et al., 2005) or possibly the SPARQL Inferencing Notation (SPIN) Knublauch et al. (2011). With the scaffolding provided through PHIDO we can also offer the ontology as an ontological framework to allow other users to implement their dialogue for health-based patient discussion. We have also demonstrated the use of ontologies for process-

based use cases, outside of dialogue.

### 6.3 FUTURE DIRECTION AND LIMITATIONS

While we have the blueprint and technology to develop and deploy (pilot test) a fully automated prototype of a conversational agent for the HPV vaccine, there are many research and development directions to pursue.

With regards to the user perception, we reported that some analyses were minimal or of no significance, and this could be attributed to a small sample size. Studies of our scale that involve participants to be physically in a live environment (i.e. as opposed to taking a survey online) require immense time and resources that were limited in our effort. Our results from the extended usability study focused on young adult students. It remains to be seen if we would attain similar perception or impact if experimented with parents, even though our early trial with parents showed some interest by that demographic. A future experiment will need to include adults to show more evidence of usefulness for the system. In a later passage, we talk about tailoring, but for different audiences we need to determine appropriate utterances for specific users. For example, using slang words for young adults – for humor or connection with the user. Our early simulation assessment, none of the users were reported to be vaccine hesitant. It would be of interest to see how vaccine hesitant parents respond to a conversational agent for HPV vaccine.

**DOMAIN LAYER:** From the domain knowledge base, VISO could benefit from linked knowledge sources to expand the scope of information available for the user. For example, the HPV vaccine is a cancer preventive measure, and some of the questions posed by the users in the simulation were cancer-related. VISO could benefit from another ontology pertaining to cancers caused by HPV infection. We envision the possibility of cancer ontology for patient knowledge that have patient-level information that are of concerns for everyday consumers. This would contrast with some of

the professional level ontologies that are more “scientific”.

**APPLICATION LAYER:** In our Wizard of OZ simulations, we experimented with two types of audiences - young adults and parents with a child under 18. One of the opportunities that our ontology for dialogue could provide is tailoring the dialogue for specific audiences. This may require an additional step in capturing contextual information from the user. The early drafts of the dialogue script included small talk and inquiries about the user before it was edited to a bare minimum. There are a few future steps to consider to make this tool more dynamic for individual users - develop a user context ontology, expand the dialogue exchange to probe the user for information, and populate the ontology with appropriate utterances specific to user, and provide rules and reasoning to coordinate tailored dialogue.

Our work with VEO and PHIDO introduces opportunities to associate emotions with utterances. For each utterance class type, we could align it with the visualized emotion to supplement the speech of the machine.

**INTERFACE LAYER:** COO-PHIDO discerns the various participant utterances based on simple similarity. We assume that in the future depending on the interaction with the users, the utterances may become more sophisticated if this work is expanded. This would mean that basic string similarity to match the utterance of the user with the string text associated with an utterance may not be sufficient. Utilizing advanced approaches like machine learning to match utterance strings with a Utterance class might be one option to consider.

In regards to question answering component, while the evaluators appear to be satisfied with most of the responses to questions, we may need to show portability. An issue with ontology-based QA is that tightly coupled systems that perform well with tested ontologies, but there is a requirement to hack the system to make it usable for another ontology knowledge base (Damjanovic et al.,



2011). We need to assess its portability by evaluating system with ontologies other than VISO-HPV. Another evaluation aspect was the presentation of the answers. About half of the responses were acceptable to the evaluators, and thus, we may need to develop ways to generate a natural language response from triples that sound “natural”. Also, some of the responses for the questions were too verbose, and this may be due to the ranking and selection criteria and require us to develop an alternative method for composing a response. While we have experimented and utilized natural language generation as an ontology evaluation tool, we could utilize the Hootation API library or components of it, like SimpleNLG to provide better automated natural language. For the most part, the question answering component is a subordinate function of the overall dialogue interaction, but we have to consider having the answer to a question spoken verbally versus reading it may elicit different outcomes for presentation and whether the answer is a proper response to the question. Inter-rater agreement was performed for 3 undergraduate pre-med students and 1 graduated students with a biology degree. The agreement of the results may be sufficient to educated young adults, like the sample in the extended simulation. It remains to be seen whether the agreement would be similar with adults with children and health providers who work directly with patients. A future direction would be to incorporate the feedback for these two groups to get a comprehensive evaluation of the question answering system.

One area where we did not explore was open sourced, offline speech recognition. We relied on commercial offerings to perform some offline speech recognition on the iOS platform, but due to the closed nature of the Apple iOS system there were limits on what we could do on their platform. In several cases, from the dialogue that we collected there were apparent transcribing errors. Furthermore, offline speech recognition we believe could be of benefit in the health care domain where, if fully utilized, can protect the utterances of the user where none of their information is sent to another cloud server. Further research in this area could also have expanded impact beyond our conversational agents.

Finally, performance was a perceived issue during the simulations. Our assumption is that a fully automated version may be faster than a touch typist or copy and paste operator. However, our next goal is to deploy and test the software engine with potential users and assess its performance for possible use in a clinical environment.

WITH the work we have accomplished in the last few years, we have laid a solid foundation to explore various informatics-related topics but more importantly this work has the potential to be portable to other health care subdomains outside of the vaccine area.



Sample Ontologies from National Center  
for Biomedical Ontology Bioportal

**Table A.1:** Popular ontologies from NCBO on September 2015.

Common Terminology Criteria for Adverse Events (CTCAE)	Basic Formal Ontology (BFO)
EDAM bioinformatics operations, data types, formats, identifiers and topics (EDAM)	Biomedical Resource Ontology (BRO)
Radiology Lexicon (RADLEX)	Ontology of Clinical Research (OCRE)
Experimental Factor Ontology (EFO)	BioMedBridges Diabetes Ontology (DIAB)
Content Archive Resource Exchange Lexicon (CARELEX)	Chemical Information Ontology (CHEMINF)
BioAssay Ontology (BAO)	Ontology for General Medical Science (OGMS)
Ontology for Biomedical Investigations (OBI)	Clinical Trials Ontology (CTO)
Brain Region & Cell Type terminology (BRCT)	Allen Brain Atlas (ABA) Adult Mouse Brain Ontology (ABA-AMB)
Cell Ontology (CL)	BioModels Ontology (BIOMODELS)
Robert Hoehndorf Version of MeSH (RH-MESH)	Microbial Culture Collection Vocabulary (MCCV)
Uber Anatomy Ontology (UBERON)	Biomedical Informatics Research Network Project Lexicon (BIRNLEX)
The Drug Ontology (DRON)	Cigarette Smoke Exposure Ontology (CSEO)
Orphanet Rare Disease Ontology (ORDO)	Translational Medicine Ontology (TMO)
Symptom Ontology (SYMP)	Mathematical Modelling Ontology (MAMO)
NanoParticle Ontology (NPO)	Pathogenic Disease Ontology (PDO)
Software Ontology (SWO)	Bilingual Ontology of Alzheimer's Disease and Related Diseases (ONTOAD)
International Classification of Functioning, Disability and Health (ICF)	Anatomic Pathology Lexicon (PATHLEX)
Statistics Ontology (STATO)	The Drug-Drug Interactions Ontology (DINTO)
Semanticscience Integrated Ontology (SIO)	Plant Experimental Assay Ontology (PEAO)
Semantic Web for Earth and Environment Technology Ontology (SWEET)	Regional Healthcare System Interoperability and Information Exchange Measurement Ontology (HEIO)
Multiple sclerosis ontology (MSO)	Time Event Ontology (TEO)
Neuroscience Information Framework (NIF) Standard Ontology (NIFSTD)	Dermatology Lexicon (DERMLEX)
Cell Line Ontology (CLO)	WikiPathways (WIKIPATHWAYS)
Radiation Oncology Ontology (ROO)	Microarray and Gene Expression Data Ontology (MO)
Information Artifact Ontology (IAO)	VIVO Ontology for Researcher Discovery (VIVO)
Kinetic Simulation Algorithm Ontology (KISAO)	Sage Bionetworks Synapse Ontology (SYN)
Genomic Clinical Decision Support Ontology (GENE-CDS)	Suggested Ontology for Pharmacogenomics (SOPHARM)
Metagenome and Microbes Environmental Ontology (MEO)	Ontology for Biobanking (OBIB)
Ortholog Ontology (ORTHO)	Cell Behavior Ontology (CBO)
Autism Spectrum Disorder Phenotype Ontology (ASDPTO)	
Breast Cancer Grading Ontology (BCGO)	
Infectious Disease Ontology (IDO)	
Drug Interaction Knowledge Base Ontology (DIKB)	
DICOM Controlled Terminology (DCM)	
Galen Ontology (GALEN)	
Vaccine Ontology (VO)	
Alzheimer's disease ontology (ADO)	



## Vaccine Information Statement Ontology

**Table B.1:** Vaccine Information Statement Ontology Class Table

Meta-level Class	Description	Example Subclasses or <i>Instances (in italics)</i>
Vaccine	A class description to categorize vaccines documented by the CDC's VIS documentation	AlternateVaccine
Target	Provides class specification for virus, bacteria, diseases, etc. that are prevented by Vaccine Disease	Virus, SeriousDisease, Bacteria
People	Categorizes various types of patients or groups of individuals impacted by vaccines, diseases, reactions, and other health conditions	PeopleWithCondition, Infants, PeopleWithIllness, PeopleWithModerateIllness, PeopleOfAge, Children, Adults
Source	Used to describe an origin of a vaccine's target (bacteria, disease). Can be reused in relations with other classes.	<i>bacteria, Hepatitis B virus</i>
Channel	Class of medium of travel for vaccine targets.	PeopleChannel, ObjectChannel, PeopleWithConditionChannel, DermalChannel, HumanActivityChannel, FluidChannel
Cause	For a description of a condition as a result of a vaccine target. E.g. infection, coughing spells.	<i>ear infection, long-term illness, coughing spells, pneumonia</i>
Location	Type to categorize location, specifically area of the body, affected by a health condition or reaction	FacialLocation, ThroatLocation, ArmLocation
Probability	Classification for types of probabilities enumerated in VIS documents	QualitativeProbability, QuantitativeProbability, ProbabilityInCases
Outcome	Types of effects resulting from causation, reactions, or chain of outcomes.	RareOutcome, AdultOutcome, ChildOutcome, FatalOutcome
Duration	Used for various types of descriptions for qualitative length of time for effects of health conditions or signs of conditions.	DurationInMinutes, DurationInWeeks
Substance	Classification of kinds of substance for vaccines and possibly other class groupings.	LiquidSubstance, GaseousSubstance, SolidSubstance
Combination	For various artifacts that interact with vaccines.	SafeCombination, DangerousCombination
Method	Groups and classifies inoculation methods for vaccines.	InhaleMethod, InjectionMethod, OralMethod
NumberOfDoses	Enumerates the maximum number of doses for vaccine.	OptionalNumberOfDoses
Dosage	Designates the types of dose or the dose interval. E.g. 1st Dose, 3rd Dose.	OptionalDose, DoseIntervals
Component	Categorizes types of elements of a vaccine.	ViralComponent, NoninfectiousViralComponent
Age	Enumerates the type of quantitative classification of age ? years, months, weeks, etc.	AgeInMonths, AgeInYears
Date	Enumerates the types of quantitative classification of date ? days, months, weeks, etc.	DateInYears, DateInMonths, DateInDays
Occurrence	Classification of types of events	VaccinatedOccurrence, TimedOccurrence
Action	Types of patient recourse in response to reactions or actions required vaccine patients before inoculation.	InactiveAction, ActiveAction, EmergencyAction
Allergen	For classes of substances leading to an allergic reaction.	VaccineAllergen, VaccineComponentAllergen
Reaction	Used to categorizes types of reactions that may result from vaccination.	MildReaction, TemporaryReaction, SeriousReaction, ModerateReaction, AdultReaction
Sign	Class of indicators for vaccine reaction effects.	<i>fast heartbeat, crying, stomach pain, hives</i>

**Table B.2:** Vaccine Information Statement Ontology's object properties

Domain	properties	Range
Target, Cause, Reaction	affects	People, Location
Dosage	after	Date
PeopleWithAllergicReaction, AllergicReaction	allergic to	Allergen
Dosage	before	Occurrence
PeopleBornFrom	born from	Date
Target	causes	Cause
Vaccine	contains	Component
Vaccine	discouraged for	People
NumberOfDoses	for	People
Cause, Outcome, Reaction	happens	Probability
Vaccine	has alternate	Alternate Vaccine
Vaccine	has dosage	Dosage
Vaccine	has number of	NumberOfDoses
Reaction	has signs	Sign
Vaccine	is a substance of	Substance
Vaccine	is safe for	People
Vaccine	is safe with	Combination
Vaccine	is taken	Method
Target, Reaction, Sign, ObjectChannel	lasts	Duration
Target, Cause, Outcome, Reaction	leads	Outcome
Cause, Reaction	located	Location
Vaccine	may cause	Reaction
Reaction	need	Action
Reaction, Sign	occurs after	Date, Occurrence
PeopleOfAge	of age	Age
Target	originates	Source
Vaccine	prevents	Target
Vaccine	protects	People
Target	spreads through	Channel
People	take	Action
Dosage	taken at	Age
Cause	to	People
HumanActivityChannel	with	People

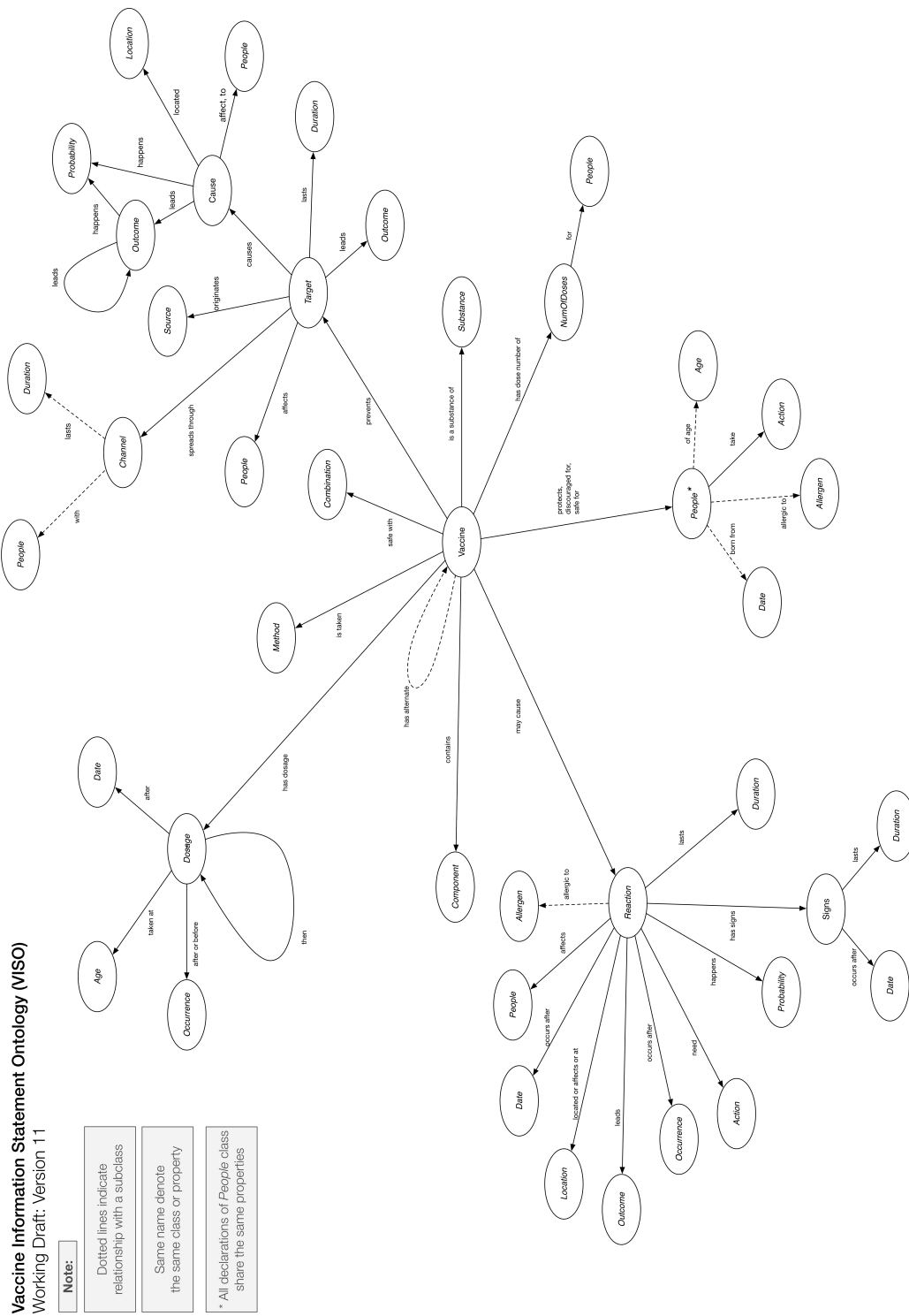


Figure B.1: TBox level structure of VISO.





# Definition of emotions from the Visualized Emotion Ontology

**Table C.1:** Definition of positive emotions.

Definition	OWL2 Axiom
positive is a valenced reaction (to “something”)	$\text{Positive\_Emotion} \sqsubseteq \text{Emotion}$
pleased is being positive about a consequence (of an event)	$\text{Pleased} \sqsubseteq \text{Positive\_Emotion} \sqcap (\exists \text{concernsConsequence.Consequence})$
hope is being pleased about a prospective consequence (of an event)	$\text{Hope} \sqsubseteq \text{Pleased} \sqcap (\exists \text{concernsConsequence.Prospective\_Consequence})$
joy is being pleased about an actual consequence (of an event)	$\text{Joy} \sqsubseteq \text{Pleased} \sqcap (\exists \text{concernsConsequence.Actual\_Consequence})$
satisfaction is joy about the confirmation of a prospective desirable consequence	$\text{Satisfaction} \sqsubseteq \text{Joy} \sqcap (\exists \text{concernsConsequence.Confirmed\_Desirable\_Consequence})$
relief is joy about the disconfirmation of a prospective undesirable consequence	$\text{Relief} \sqsubseteq \text{Joy} \sqcap (\exists \text{concernsConsequence.Disconfirmed\_Undesirable\_Consequence})$
happy-for is joy about a consequence (of an event) presumed to be desirable for someone else	$\text{Happy\_For} \sqsubseteq \text{Joy} \sqcap (\exists \text{concernsConsequence.Consequence\_Desirable\_For\_Other})$
gloating is joy about a consequence (of an event) presumed to be undesirable for someone else	$\text{Gloating} \sqsubseteq \text{Joy} \sqcap (\exists \text{concernsConsequence.Consequence\_Undesirable\_For\_Other})$
surprise is joy or distress about the disconfirmation of a prospective undesirable consequence	$\text{Surprise} \sqsubseteq \text{Joy} \sqcap (\exists \text{concernsConsequence.Disconfirmed\_Consequence})$ $\text{Surprise} \sqsubseteq \text{Distress} \sqcap (\exists \text{concernsConsequence.Disconfirmed\_Consequence})$
approving is being positive about an action (of an agent)	$\text{Approving} \sqsubseteq \text{Positive\_Emotion} \sqcap (\exists \text{concernsActionsOf.Action})$
pride is approving of one’s own action	$\text{Pride} \sqsubseteq \text{Approving} \sqcap (\exists \text{concernsActionsOf.Action\_Of\_Self\_Agent})$
admiration is approving of someone else’s action	$\text{Admiration} \sqsubseteq \text{Approving} \sqcap (\exists \text{concernsActionsOf.Action\_Of\_Other\_Agent})$
gratification is pride about an action and joy about a related consequence	$\text{Gratification} \sqsubseteq \text{Pride} \text{ Gratification} \sqsubseteq \text{Joy} \text{ (n.b.)}$
gratitude is admiration about an action and joy about a related consequence	$\text{Gratitude} \sqsubseteq \text{Admiration} \text{ Gratitude} \sqsubseteq \text{Joy}$
liking is being positive about an aspect (of an object)	$\text{Liking} \sqsubseteq \text{Positive\_Emotion} \sqcap (\exists \text{concernsAspect.Aspect})$
love is liking a familiar aspect (of an object)	$\text{Love} \sqsubseteq \text{Liking} \sqcap (\exists \text{concernsAspect.Familiar\_Aspect})$
interest is liking an unfamiliar aspect (of an object)	$\text{Interest} \sqsubseteq \text{Liking} \sqcap (\exists \text{concernsAspect.Unfamiliar\_Aspect})$

**Table C.2:** Definition of negative emotions

Definition	OWL2 Axiom
negative is a valenced reaction (to “something”)	Negative_Emotion $\sqsubseteq$ Emotion
displeased is being negative about a consequence (of an event)	Displeased $\sqsubseteq$ Negative_Emotion $\sqcap (\exists \text{concernsConsequence.Consequence})$
fear is being displeased about a prospective consequence (of an event)	Fear $\sqsubseteq$ Displeased $\sqcap (\exists \text{concernsConsequence.Prospective_Consequence})$
distress is being displeased about an actual consequence (of an event)	Distress $\sqsubseteq$ Displeased $\sqcap (\exists \text{concernsConsequence.Actual_Consequence})$
fears-confirmed is distress about the confirmation of a prospective undesirable consequence	Fears_Confirmed $\sqsubseteq$ Distress $\sqcap (\exists \text{concernsConsequence.Confirmed_Undesirable_Consequence})$
disappointment is distress about the disconfirmation of a prospective desirable consequence	Disappointment $\sqsubseteq$ Distress $\sqcap (\exists \text{concernsConsequence.Disconfirmed_Desirable_Consequence})$
resentment is distress about a consequence (of an event) presumed to be desirable for someone else	Resentment $\sqsubseteq$ Distress $\sqcap (\exists \text{concernsConsequence.Consequence_Desirable_For_Other})$
pity is distress about a consequence (of an event) presumed to be undesirable for someone else	Pity $\sqsubseteq$ Distress $\sqcap (\exists \text{concernsConsequence.Consequence_Undesirable_For_Other})$
surprise is joy or distress about the disconfirmation of a prospective undesirable consequence	Surprise $\sqsubseteq$ Joy $\sqcap (\exists \text{concernsConsequence.Disconfirmed_Consequence})$ Surprise $\sqsubseteq$ Distress $\sqcap (\exists \text{concernsConsequence.Disconfirmed_Consequence})$
disapproving is being negative about an action (of an agent)	Disapproving $\sqsubseteq$ Negative_Emotion $\sqcap (\exists \text{concernsActionsOf.Action})$
shame is disapproving of one’s own action	Shame $\sqsubseteq$ Disapproving $\sqcap (\exists \text{concernsActionsOf.Action_Of_Self_Agent})$
reproach is disapproving of someone else’s action	Reproach $\sqsubseteq$ Disapproving $\sqcap (\exists \text{concernsActionsOf.Action_Of_Other_Agent})$
remorse is shame about an action and distress about a related consequence	Remorse $\sqsubseteq$ Shame Remorse $\sqsubseteq$ Distress
anger is reproach about an action and distress about a related consequence	Anger $\sqsubseteq$ Reproach Anger $\sqsubseteq$ Distress
disliking is being negative about an aspect (of an object)	Disliking $\sqsubseteq$ Negative_Emotion $\sqcap (\exists \text{concernsAspect.Aspect})$
hate is disliking a familiar aspect (of an object)	Hate $\sqsubseteq$ Disliking $\sqcap (\exists \text{concernsAspect.Familiar_Aspect})$
disgust is disliking an unfamiliar aspect (of an object)	Disgust $\sqsubseteq$ Disliking $\sqcap (\exists \text{concernsAspect.Unfamiliar_Aspect})$



Frankenstein Ontology

Question-Answering for User-Centric  
System

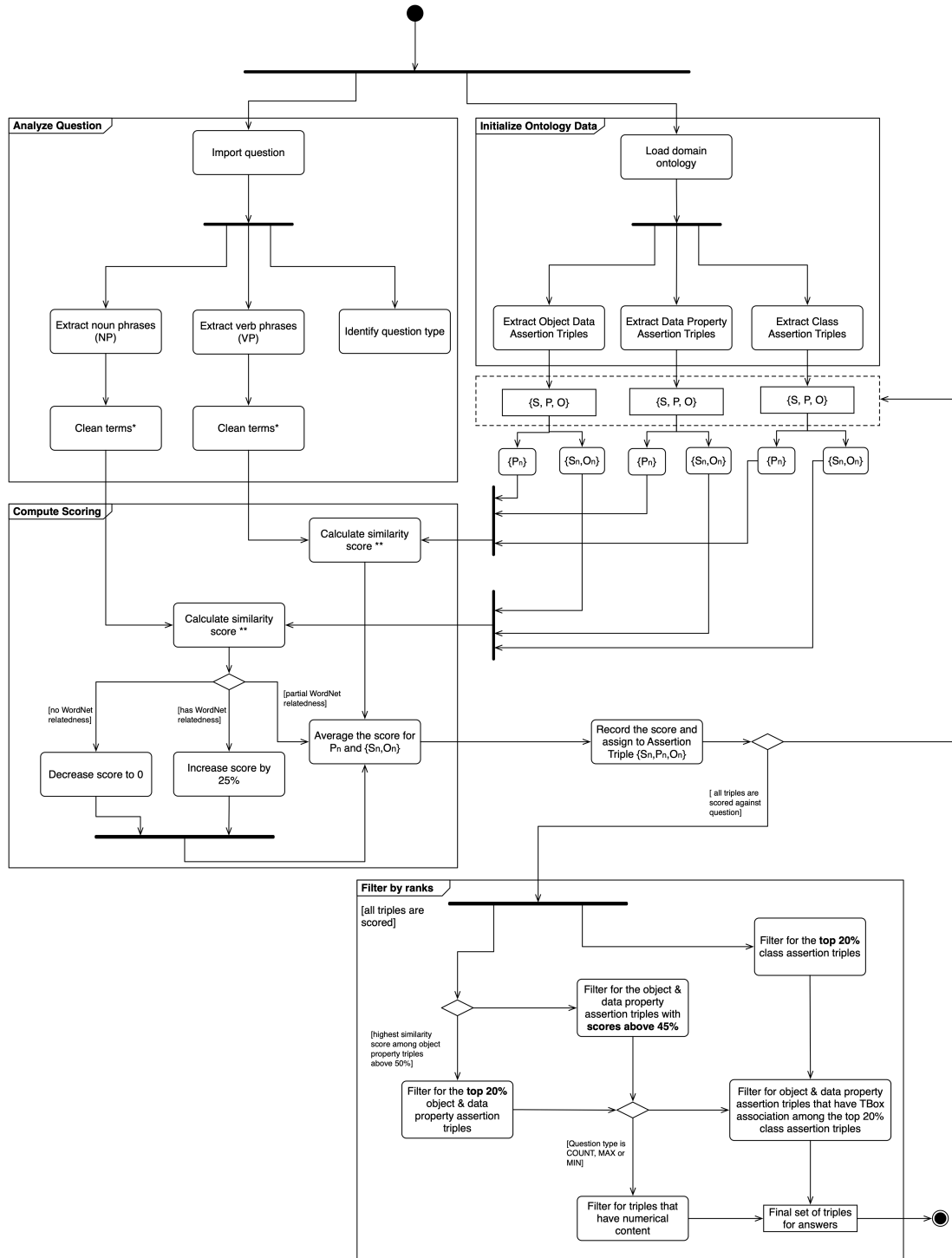


Figure D.1: Process diagram of Frankenstein Ontology Question-Answering for User-Centric System

# E

## Usability Variables

**Table E.1:** Participant demographic variables

Variable ID	SPSS Code	Measurement/Value
<b>DEMOGRAPHIC BLOCK</b>		
Participant	participants	Scale
Participant's self identified gender	gender	Nominal {m, f}
Participant's HPV vaccination status	hpv_vax_status	Nominal {1=yes,2=no,3=don't know}
Participant's race/ethnicity	ethnicity	Nominal {white, black, hispanic, asian}
Participant's age	age	Scale
Participant's college classification	Class	Nominal {freshman, sophomore, junior, senior}
Income	Income	Nominal {a= less than 5K
		b= 5001-10K;
		c= 10,001 - 20K;
		d= 20,001 - 30K;
		e= 30,001 - 40K;
		f= 40,001- 50K;
		g= 50K+
Parent's income	ParentIncome	Nominal {1= less than 25,000; 2=25K - 34K;
		3=35K-49K;
		4=50K-74,999;
		5=75K-99,999;
		6=100K-149,999;
		7= 150K+}
<b>EXPERIENCE BLOCK</b>		
Technical difficulties experienced	Technical_difficulties	Nominal {0 = no, 1= yes}

**Table E.2:** System Usability Survey variables

Variable ID	SPSS Code	Measurement/Value
<b>SYSTEM USABILITY SCALE</b>		
SUS computed score	SUS_Score	Scale
SUS school grade	SUS_Grade	Ordinal {A, B, C, D, F}
SUS adjective	SUS_Adjective	Ordinal {Worst imaginable, Awful, Poor, OK, Good, Excellent, Best Imaginable}

**Table E.3: Speech User Interface Service Quality survey**

Variable ID	SPSS Code	Measurement/Value
<b>USER GOAL ORIENTATION BLOCK</b>		
Beverly made me feel like I was in control	UGO <sub>1</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
Beverly gave me a good feeling about being a health consumer	UGO <sub>3</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
I could find what I needed without any difficulty	UGO <sub>5</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
Beverly would help me to be productive	UGO <sub>10</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
I could trust Beverly to work correctly	UGO <sub>12</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
I would be likely to use Beverly again	UGO <sub>13</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
I feel confident using Beverly	UGO <sub>17</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
The quality of Beverly made me want to remain a health consumer	UGO <sub>19</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
<b>CUSTOMER SERVICE BEHAVIOR BLOCK</b>		
Beverly used terms I am familiar with	CSB <sub>4</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
Beverly used everyday words	CSB <sub>6</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
Beverly was organized and logical	CSB <sub>7</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
Beverly spoke at a pace that was easy to follow	CSB <sub>9</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
Beverly seemed polite	CSB <sub>11</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
Beverly seemed courteous	CSB <sub>21</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
Beverly seemed friendly	CSB <sub>23</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
Beverly seemed profession in its speaking style	CSB <sub>25</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
<b>SPEECH CHARACTERISTICS BLOCK</b>		
Beverly's voice was pleasant	SC <sub>14</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
Beverly voice sounded like people I hear on media devices	SC <sub>16</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
Beverly's voice sounded like a regular person	SC <sub>18</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
Beverly's voice sounded natural	SC <sub>20</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
Beverly's voice sounded enthusiastic and full of energy	SC <sub>24</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
<b>VERBOSITY BLOCK</b>		
The messages were repetitive	V <sub>2</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
Beverly gave me more details that I needed	V <sub>8</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
Beverly was too talkative	V <sub>15</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
I felt like I had to wait too long for Beverly to stop talking so I could respond	V <sub>22</sub>	Ordinal {1=Strongly disagree – 7=Strongly Agree}
<b>SUISQ Scoring BLOCK</b>		
User Goal Orientation Score	User_Goal_Orientation	Ordinal {0-7}
Customer Service Behavior Score	Customer_Service_Behavior	Ordinal {0-7}
Speech Characteristic Score	Speech_Characteristics	Ordinal {0-7}
Verbosity Score	Verbosity	Ordinal {0-7}
Verbosity Reversed (to calculate SUIQ)	Verbosity_Reverse	Ordinal {0-7; 8- Verbosity_R}
SUISQ Overall Score	Overall_SUISQ	Ordinal {0-7; mean of factors replacing Verbosity with Verbosity_Reverse}



**Table E.4:** CHIAS health belief model (HBM)

Variable ID	SPSS Code	Measurement/Value
<b>Harms Factor BLOCK</b>		
The HPV vaccine might cause short term problems, like fever or discomfort (Harms)	H_1	Ordinal {1=Strongly disagree – 4=Strongly agree}
The HPV vaccine is being pushed to make money for drug companies (Harms)	H_2	Ordinal {1=Strongly disagree – 4=Strongly agree}
I think the HPV vaccine might cause health problems in the future (Harms)	H_3	Ordinal {1=Strongly disagree – 4=Strongly agree}
I think the HPV vaccine is unsafe (Harms)	H_5	Ordinal {1=Strongly disagree – 4=Strongly agree}
The HPV vaccine is so new that I want to wait a while before deciding if I should get it (Harms)	H_15	Ordinal {1=Strongly disagree – 4=Strongly agree}
<b>Barrier Factor BLOCK</b>		
It would be hard to find a provider or clinic where I could afford the HPV vaccine (Barriers)	B_7	Ordinal {1=Strongly disagree – 4=Strongly agree}
It would be hard to find a provider or clinic that would be easy to get to for getting vaccinated against HPV (Barriers)	B_8	Ordinal {1=Strongly disagree – 4=Strongly agree}
It would be hard to find a provider or clinic that has the HPV vaccine available (Barriers)	B_9	Ordinal {1=Strongly disagree – 4=Strongly agree}
I am concerned that the HPV vaccine costs more than I can pay (Barriers)	B_10	Ordinal {1=Strongly disagree – 4=Strongly agree}
It would be hard to find a provider or clinic where I don't have to wait a long time to get an appointment to be vaccinated (Barriers)	B_11	Ordinal {1=Strongly disagree – 4=Strongly agree}
<b>Effectiveness Factor BLOCK</b>		
How effective do you think the HPV vaccine is in preventing genital warts (Effectiveness)	E_12	Ordinal {1=Slightly effective – 4=Extremely effective}
How effective do you think the HPV vaccine is in preventing cervical cancer (Effectiveness)	E_13	Ordinal {1=Slightly effective – 4=Extremely effective}
<b>Uncertainty Factor BLOCK</b>		
I have enough information about the HPV vaccine to decide whether to get it (Uncertainty)	U_14	Ordinal {1=Strongly disagree – 4=Strongly agree}
My friends are getting the HPV vaccine (Uncertainty)	U_16	Ordinal {1=Strongly disagree – 4=Strongly agree}
<b>Risk Denial (young adults) BLOCK</b>		
I think that getting the HPV vaccine makes it more likely for someone to have sex (Risk Denial)	RD_4	Ordinal {1=Strongly disagree – 4=Strongly agree}
I think I am too young to get a vaccine for a sexually transmitted infection like HPV (Risk Denial)	RD_6	Ordinal {1=Strongly disagree – 4=Strongly agree}
HPV vaccination is not really necessary because Pap smears can be done to make sure cervical cancer doesn't develop (Risk Denial)	RD_17	Ordinal {1=Strongly disagree – 4=Strongly agree}

# References

- Adams, W. G., Fuhlbrigge, A. L., Miller, C. W., Panek, C. G., Gi, Y., Loane, K. C., Madden, N. E., Plunkett, A. M., & Friedman, R. H. (2003). Tlc-asthma: an integrated information system for patient-centered monitoring, case management, and point-of-care decision support. In *AMIA Annual Symposium Proceedings*, volume 2003 (pp.1): American Medical Informatics Association.
- Afify, Y. M., Badr, N. L., Moawad, I. F., & Tolba, M. F. (2017). Evaluation of cloud service ontologies. In *2017 Eighth International Conference on Intelligent Computing and Information Systems (ICICIS)*(pp. 144–153): IEEE.
- Agre, P. E. & Chapman, D. (1987). Pengi: An implementation of a theory of activity. In *AAAI*, volume 87 (pp. 286–272).
- Ahken, S., Fleming, N., Dumont, T., & Black, A. (2015). Hpv awareness in higher-risk young women: The need for a targeted hpv catch-up vaccination program. *J Obstet Gynaecol Can*, 37(2), 122–128.

- Al-Naggar, R. A., Al-Jashamy, K., & Chen, R. (2010). Perceptions and opinions regarding human papilloma virus vaccination among young women in malaysia. *Asian Pac J Cancer Prev*, 11(6), 1515–1521.
- Alani, H. & Brewster, C. (2006). Metrics for ranking ontologies. In *Evaluating Ontologies for the Web Workshop (EON2006)*.
- Allemang, D. & Hendler, J. (2011). *Semantic web for the working ontologist: effective modeling in RDFS and OWL*. Elsevier.
- Allen, J., Ferguson, G., Blaylock, N., Byron, D., Chambers, N., Dzikovska, M., Galescu, L., & Swift, M. (2006). Chester: towards a personal medication advisor. *Journal of biomedical informatics*, 39(5), 500–513.
- Allen, J. D., de Jesus, M., Mars, D., Tom, L., Cloutier, L., & Shelton, R. C. (2011). Decision-making about the hpv vaccine among ethnically diverse parents: implications for health communications. *Journal of oncology*, 2012.
- Allen, J. D., Mohllajee, A. P., Shelton, R. C., Othus, M. K., Fontenot, H. B., & Hanna, R. (2009). Stage of adoption of the human papillomavirus vaccine among college women. *Preventive medicine*, 48(5), 420–425.
- Allnatt, R., Engelhardt, J., Gao, J., & Mariano, A. (2001). An evaluation of the functionality and acceptability of the voice prescription label. *Journal of Visual Impairment & Blindness (JVIB)*, 95(11).

- Almeida, M. B. (2009). A proposal to evaluate ontology content. *Applied Ontology*, 4(3), 245–265.
- Amith, M., Cohen, T., Cunningham, R., Savas, L. S., Smith, N., Cuccaro, P., Gabay, E., Boom, J., Schvaneveldt, R., & Tao, C. (In Reviewa). Mining hpv vaccine knowledge structures of young adults from reddit using distributional semantics and pathfinder networks. *Cancer Control Journal*.
- Amith, M., Cunningham, R., Savas, L. S., Boom, J., Schvaneveldt, R., Tao, C., & Cohen, T. (2017a). Using Pathfinder networks to discover alignment between expert and consumer conceptual knowledge from online vaccine content. *Journal of Biomedical Informatics*, 74, 33–45.
- Amith, M., Gong, Y., Cunningham, R., Boom, J., & Tao, C. (2015). Developing VISO: Vaccine Information Statement Ontology for patient education. *Journal of Biomedical Semantics*, 6(1), 23.
- Amith, M., He, Z., Bian, J., Lossio-Ventura, J. A., & Tao, C. (2018a). Assessing the practice of biomedical ontology evaluation: Gaps and opportunities. *Journal of Biomedical Informatics*, 80, 1–13.
- Amith, M., Manion, F., Liang, C., Harris, M., Wang, D., He, Y., & Tao, C. (2018b). Ontokeeper: Semiotic-driven ontology evaluation tool for biomedical ontologists. In *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (pp. 1614–1617).

- Amith, M., Manion, F., Liang, C., Harris, M., Wang, D., He, Y., & Tao, C. (2019a). Architecture and usability of ontokeeper, an ontology evaluation tool. *BMC Medical Informatics and Decision Making*, 225, 432–436.
- Amith, M., Manion, F. J., Harris, M. R., Zhang, Y., Xu, H., & Tao, C. (2017b). Expressing Biomedical Ontologies in Natural Language for Expert Evaluation. *Studies in Health Technology and Informatics*, (pp. 838–842).
- Amith, M., Roberts, K., & Tao, C. (In Reviewb). Conceiving an application ontology to model patient human papillomavirus vaccine counseling for dialogue management. *BMC Bioinformatics*.
- Amith, M. & Tao, C. (2015). A web application towards semiotic-based evaluation of biomedical ontologies. In *Proceedings of International Workshop on Biomedical Data Mining, Modeling, and Semantic Integration: A Promising Approach to Solving Unmet Medical Needs (BDM2I 2015)*.
- Amith, M. & Tao, C. (2017). Modulated evaluation metrics for drug-based ontologies. *Journal of Biomedical Semantics*, 8(1), 17.
- Amith, M. & Tao, C. (2018). Representing vaccine misinformation using ontologies. *Journal of biomedical semantics*, 9(1), 22.

- Amith, M., Zhu, A., Cunningham, R., Lin, R., Savas, L., Shay, L., Chen, Y., Gong, Y., Boom, J., Roberts, K., & Tao, C. (2019b). Early Usability Assessment of a Conversational Agent for HPV Vaccination. *Studies in Health Technology and Informatics*, (pp. 17–23).
- Androutsopoulos, I., Ritchie, G. D., & Thanisch, P. (1995). Natural language interfaces to databases—an introduction. *Natural language engineering*, 1(01), 29–81.
- Angeli, G., Premkumar, M. J. J., & Manning, C. D. (2015). Leveraging linguistic structure for open domain information extraction. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, volume 1 (pp. 344–354).
- Anhang Price, R., Tiro, J. A., Saraiya, M., Meissner, H., & Breen, N. (2011). Use of human papillomavirus vaccines among young adult women in the united states: an analysis of the 2008 national health interview survey. *Cancer*, 117(24), 5560–5568.
- Apache Software Foundation (2009). Apache jena. <https://jena.apache.org/>.
- Aruna, T., Saranya, K., & Bhandari, C. (2011). A survey on ontology evaluation tools. In *Process Automation, Control and Computing (PACC), 2011 International Conference on* (pp. 1–5).: IEEE.
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., et al. (2000). Gene ontology: tool for the unification of biology. *Nature genetics*, 25(1), 25–29.

- Askelson, N. M., Campo, S., Lowe, J. B., Smith, S., Dennis, L. K., & Andsager, J. (2010). Using the theory of planned behavior to predict mothers' intentions to vaccinate their daughters against hpv. *The Journal of School Nursing*, 26(3), 194–202.
- Atkinson, W., Wolfe, C., & Hamborsky, J. (2012). *Centers for Disease Control and Prevention. Epidemiology and Prevention of Vaccine-Preventable Diseases*. Washington DC: Public Health Foundation, 12th edition.
- Autayeu, A. (2016). extJWNL. <http://extjwnl.sourceforge.net/>.
- Bair, R. M., Mays, R. M., Sturm, L. A., & Zimet, G. D. (2008). Acceptability of the human papillomavirus vaccine among latina mothers. *Journal of pediatric and adolescent gynecology*, 21(6), 329–334.
- Bakogianni, G. D., Nikolakopoulos, K. M., & Nikolakopoulou, N. M. (2010). Hpv vaccine acceptance among female greek students. *International journal of adolescent medicine and health*, 22(2), 271–274.
- Bangor, A. (2009). Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *Journal of Usability Studies*, 4(3), 10.
- Banko, M., Cafarella, M. J., Soderland, S., Broadhead, M., & Etzioni, O. (2007). Open information extraction for the web. In *IJCAI*, volume 7 (pp. 2670–2676).
- Bartholomew, L. K., Parcel, G. S., Kok, G., Gottlieb, N. H., et al. (2011). *Planning health promotion programs: an intervention mapping approach*. John Wiley & Sons.

- Bast, H. & Haussmann, E. (2013). Open information extraction via contextual sentence decomposition. In *2013 IEEE Seventh International Conference on Semantic Computing* (pp. 154–159).: IEEE.
- Bastani, R., Glenn, B. A., Tsui, J., Chang, L. C., Marchand, E. J., Taylor, V. M., & Singhal, R. (2011). Understanding suboptimal human papillomavirus vaccine uptake among ethnic minority girls. *Cancer Epidemiology Biomarkers & Prevention*, 20(7), 1463–1472.
- Bean, S. J. (2011). Emerging and continuing trends in vaccine opposition website content. *Vaccine*, 29(10), 1874–1880.
- Becker, M. H. (1974). The health belief model and personal health behavior. *Health education monographs*, 2, 324–473.
- Bernard, D. M., Cooper Robbins, S. C., McCaffery, K. J., Scott, C. M., & Skinner, S. R. (2011). The domino effect: adolescent girls' response to human papillomavirus vaccination. *Med J Aust*, 194(6), 297–300.
- Berners-Lee, T., Hendler, J., Lassila, O., et al. (2001). The semantic web. *Scientific american*, 284(5), 28–37.
- Betsch, C., Renkewitz, F., Betsch, T., & Ulshöfer, C. (2010). The influence of vaccine-critical websites on perceiving vaccination risks. *Journal of Health Psychology*, 15(3), 446–455.
- Beveridge, M. & Fox, J. (2006). Automatic generation of spoken dialogue from medical plans and ontologies. *Journal of biomedical informatics*, 39(5), 482–499.



- Bickmore, T. & Giorgino, T. (2006). Health dialog systems for patients and consumers. *Journal of biomedical informatics*, 39(5), 556–571.
- Bickmore, T. W., Pfeifer, L. M., Byron, D., Forsythe, S., Henault, L. E., Jack, B. W., Silliman, R., & Paasche-Orlow, M. K. (2010). Usability of conversational agents by patients with inadequate health literacy: evidence from two clinical trials. *Journal of health communication*, 15 Suppl 2, 197.
- Bickmore, T. W., Schulman, D., & Sidner, C. (2013). Automated Interventions for Multiple Health Behaviors Using Conversational Agents. *Patient education and counseling*, 92(2), 142–148.
- Bickmore, T. W., Schulman, D., & Sidner, C. L. (2011). A reusable framework for health counseling dialogue systems based on a behavioral medicine ontology. *Journal of Biomedical Informatics*, 44(2), 183–197.
- Blake, K. D., Ottenbacher, A. J., Rutten, L. J. F., Grady, M. A., Kobrin, S. C., Jacobson, R. M., & Hesse, B. W. (2015). Predictors of human papillomavirus awareness and knowledge in 2013: Gaps and opportunities for targeted communication strategies. *American journal of preventive medicine*, 48(4), 402–410.
- Bodenreider, O. & Stevens, R. (2006). Bio-ontologies: current trends and future directions. *Briefings in Bioinformatics*, 7(3), 256–274.

- Boehner, C. W., Howe, S. R., Bernstein, D. I., & Rosenthal, S. L. (2003). Viral sexually transmitted disease vaccine acceptability among college students. *Sexually Transmitted Diseases*, 30(10), 774–778.
- Brabin, L., Roberts, S. A., Farzaneh, F., & Kitchener, H. C. (2006). Future acceptance of adolescent human papillomavirus vaccination: a survey of parental attitudes. *Vaccine*, 24(16), 3087–3094.
- Brank, J., Grobelnik, M., & Mladenić, D. (2005). A survey of ontology evaluation techniques. In *Proceedings of the conference on data mining and data warehouses (SiKDD 2005)*.
- Brawner, B. M., Baker, J. L., Voytek, C. D., Leader, A., Cashman, R. R., Silverman, R., Peter, N., Buchner, B. J., Barnes, C. A., Jemmott, L. S., et al. (2012). The development of a culturally relevant, theoretically driven hpv prevention intervention for urban adolescent females and their parents/guardians. *Health promotion practice*, (pp. 1524839912462389).
- Brewer, N. T. & Fazekas, K. I. (2007). Predictors of hpv vaccine acceptability: a theory-informed, systematic review. *Preventive medicine*, 45(2), 107–114.
- Brewer, N. T., Gottlieb, S. L., Reiter, P. L., McRee, A.-L., Liddon, N., Markowitz, L., & Smith, J. S. (2011). Longitudinal Predictors of Human Papillomavirus Vaccine Initiation Among Adolescent Girls in a High-Risk Geographic Area:. *Sexually Transmitted Diseases*, 38(3), 197–204.
- Brooke, J. (2013). SUS: a retrospective. *Journal of usability studies*, 8(2), 29–40.

- Brooke, J. et al. (1996). Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194), 4–7.
- Brooks, R. (1986). A robust layered control system for a mobile robot. *IEEE journal on robotics and automation*, 2(1), 14–23.
- Brooks, R. A. (1991a). Intelligence without reason. In *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence(IJCAI-91)* (pp. 569–595).
- Brooks, R. A. (1991b). Intelligence without representation. *Artificial intelligence*, 47(1-3), 139–159.
- Burgess, D. C., Burgess, M. A., & Leask, J. (2006). The mmr vaccination and autism controversy in united kingdom 1998–2005: Inevitable community outrage or a failure of risk communication? *Vaccine*, 24(18), 3921–3928.
- Burke, S. C., Vail-Smith, K., White, D. M., Baker, E., & Mitchell, T. (2010). Getting vaccinated against hpv: attitudes, intentions and perceived barriers of female undergraduates. *College Student Journal*, 44(1), 55.
- Burton-Jones, A., Storey, V. C., Sugumaran, V., & Ahluwalia, P. (2005). A semiotic metrics suite for assessing the quality of ontologies. *Data & Knowledge Engineering*, 55(1), 84–102.
- Bühmann, L., Usbeck, R., & Ngomo, A.-C. N. (2015). ASSESS—Automatic Self-Assessment Using Linked Data. In *The Semantic Web-ISWC 2015* (pp. 76–89). Springer.

- Carlson, A., Betteridge, J., Kisiel, B., Settles, B., Hruschka, E. R., & Mitchell, T. M. (2010a). Toward an architecture for never-ending language learning. In *Twenty-Fourth AAAI Conference on Artificial Intelligence: Association for Computational Linguistics*.
- Carlson, A., Betteridge, J., Wang, R. C., Hruschka Jr, E. R., & Mitchell, T. M. (2010b). Coupled semi-supervised learning for information extraction. In *Proceedings of the third ACM international conference on Web search and data mining* (pp. 101–110).: Association of Computing and Machinery.
- Caskey, R., Lindau, S. T., & Alexander, G. C. (2009). Knowledge and early adoption of the hpv vaccine among girls and young women: results of a national survey. *Journal of Adolescent Health*, 45(5), 453–462.
- Castellsague, X., Munoz, N., Pitisuttithum, P., Ferris, D., Monsonego, J., Ault, K., Luna, J., Myers, E., Mallery, S., Bautista, O., et al. (2011). End-of-study safety, immunogenicity, and efficacy of quadrivalent hpv (types 6, 11, 16, 18) recombinant vaccine in adult women 24–45 years of age. *British journal of cancer*, 105(1), 28–37.
- Castle, P. E. & Scarinci, I. (2009). Should hpv vaccine be given to men? *British Medical Journal Publishing Group*, (339).
- Cates, J. R., Brewer, N. T., Fazekas, K. I., Mitchell, C. E., & Smith, J. S. (2009). Racial differences in hpv knowledge, hpv vaccine acceptability, and related beliefs among rural, southern women. *The Journal of Rural Health*, 25(1), 93–97.

Centers for Disease Control and Prevention (2008). Hpv vaccine—questions and answers.

<http://www.cdc.gov/hpv/parents/questions-answers.html>.

Centers for Disease Control and Prevention (2012). Adult vaccination coverage—united states, 2010.

*MMWR. Morbidity and mortality weekly report*, 61(4), 66.

Centers for Disease Control and Prevention (2014a). HPV vaccine resources for health professions.

<http://www.cdc.gov/vaccines/who/teens/for-hcp/hpv-resources.html>.

Centers for Disease Control and Prevention (2014b). Human papillomavirus (hpv).

<http://www.cdc.gov/hpv>.

Centers for Disease Control and Prevention (2014c). Vaccine information statement: Facts about

VISs - Vaccines - CDC. <http://www.cdc.gov/vaccines/hcp/vis/about/facts-vis.html>.

Centers for Disease Control and Prevention (2016a). Centers for disease control, genital hpv

infection-std fact sheet. <http://www.cdc.gov/STD/HPV/STDFact-HPV.htm>.

Centers for Disease Control and Prevention (2016b). Vaccine Information Statement: Diphtheria,

Tetanus, and Pertussis - Vaccines - CDC. <http://www.cdc.gov/vaccines/hcp/vis/vis-statements/dtap.html>.

Centers for Disease Control and Prevention (2016c). Vaccine Information

Statement: Haemophilus Influenzae Type b (Hib) - Vaccines - CDC.

<http://www.cdc.gov/vaccines/hcp/vis/vis-statements/hib.html>.

Centers for Disease Control and Prevention (2016d). Vaccine Information Statement: Hepatitis B -

Vaccines - CDC. <http://www.cdc.gov/vaccines/hcp/vis/vis-statements/hep-b.html>.

Centers for Disease Control and Prevention (2016e). Vaccine Information Statement: MMR

(Measles, Mumps, Rubella) - Vaccines - CDC. <http://www.cdc.gov/vaccines/hcp/vis/vis-statements/mmr.html>.

Centers for Disease Control and Prevention (2016f). Vaccine Information Statement:

Pneumococcal Conjugate - Vaccines - CDC. <http://www.cdc.gov/vaccines/hcp/vis/vis-statements/pcv13.html>.

Centers for Disease Control and Prevention (2016g). Vaccine Information Statement: Rotavirus -

Vaccines - CDC. <http://www.cdc.gov/vaccines/hcp/vis/vis-statements/rotavirus.html>.

Centers for Disease Control and Prevention (2019). Vaccine information statement: Home page -

vaccines - cdc. <http://www.cdc.gov/vaccines/hcp/vis/index.html>.

Centers for Disease Control and Prevention (CDC) (2010). Fda licensure of quadrivalent human

papillomavirus vaccine (hpv4, gardasil) for use in males and guidance from the advisory committee on immunization practices (acip). *MMWR. Morbidity and mortality weekly report*, 59(20), 630.

Centers for Disease Control and Prevention (CDC) (2011). Epidemiology and prevention of

vaccine-preventable diseases. *Washington, DC: Public Health Foundation*.

- Centers for Disease Control and Prevention (CDC) (2012). Human papillomavirus-associated cancers-united states, 2004-2008. *MMWR. Morbidity and mortality weekly report*, 61, 258.
- Centers for Disease Control and Prevention (CDC) (2013a). Cdc fact sheet: incidence, prevalence, and cost of sexually transmitted infections in the united states. cdc.
- Centers for Disease Control and Prevention (CDC) (2013b). Hpv vaccine: safe, effective, and grossly underutilized.
- Centers for Disease Control and Prevention (CDC) (2013c). Human papillomavirus vaccination coverage among adolescent girls, 2007-2012, and postlicensure vaccine safety monitoring, 2006-2013-united states. *MMWR. Morbidity and mortality weekly report*, 62(29), 591.
- Centers for Disease Control and Prevention (CDC) (2013d). National and state vaccination coverage among adolescents aged 13-17 years-united states, 2012. *MMWR. Morbidity and mortality weekly report*, 62(34), 685.
- Centers for Disease Control and Prevention (CDC) (2014). How many cancers are linked with hpv each year. <http://www.cdc.gov/cancer/hpv/statistics/cases.htm>.
- CervicalCancer.org (2014). Cervical cancer statistics. <http://www.cervicalcancer.org/statistics.html>.
- Cetto, M., Niklaus, C., Freitas, A., & Handschuh, S. (2018). Graphene: Semantically-linked propositions in open information extraction. In *Proceedings of the 27th International Conference on Computational Linguistics* (pp. 2300-2311): Association for Computational Linguistics.

- Chafe, W. L. (1982). Integration and involvement in speaking, writing, and oral literature. *Spoken and written language: Exploring orality and literacy*, (pp. 35–53).
- Chao, C., Velicer, C., Slezak, J. M., & Jacobsen, S. J. (2010). Correlates for human papillomavirus vaccination of adolescent girls and young women in a managed care organization. *American Journal of Epidemiology*, (pp. kwp365).
- Charo, R. A. (2007). Politics, parents, and prophylaxis—mandating hpv vaccination in the united states. *New England Journal of Medicine*, 356(19), 1905–1908.
- Chaturvedi, A. K., Engels, E. A., Pfeiffer, R. M., Hernandez, B. Y., Xiao, W., Kim, E., Jiang, B., Goodman, M. T., Sibug-Saber, M., Cozen, W., et al. (2011). Human papillomavirus and rising oropharyngeal cancer incidence in the united states. *Journal of Clinical Oncology*, 29(32), 4294–4301.
- Chesson, H. W., Dunne, E. F., Hariri, S., & Markowitz, L. E. (2014). The estimated lifetime probability of acquiring human papillomavirus in the united states. *Sexually transmitted diseases*, 41(11), 660–664.
- Chou, B., Krill, L. S., Horton, B. B., Barat, C. E., & Trimble, C. L. (2011). Disparities in human papillomavirus vaccine completion among vaccine initiators. *Obstetrics & Gynecology*, 118(1), 14–20.



- Chow, S.-N., Soon, R., Park, J. S., Pancharoen, C., Qiao, Y. L., Basu, P., & Ngan, H. Y. S. (2010). Knowledge, attitudes, and communication around human papillomavirus (hpv) vaccination amongst urban asian mothers and physicians. *Vaccine*, 28(22), 3809–3817.
- Christel, M. G. (1994). The role of visual fidelity in computer-based instruction. *Human-Computer Interaction*, 9(3), 183–223.
- Cimino, J. J. (1998). Desiderata for controlled medical vocabularies in the twenty-first century. *Methods of information in medicine*, 37(04/05), 394–403.
- Cimino, J. J. & Zhu, X. (2006). The practical impact of ontologies on biomedical informatics. *Yearbook of medical informatics*, 15(01), 124–135.
- Clore, G. L. & Ortony, A. (2013). Psychological construction in the OCC model of emotion. *Emotion Review*, 5(4), 335–343.
- Coe, A. B., Gatewood, S. B., Moczygemba, L. R., et al. (2012). The use of the health belief model to assess predictors of intent to receive the novel (2009) h1n1 influenza vaccine. *Innovations in pharmacy*, 3(2), 1.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1), 37–46.
- Cohen, M. H., Cohen, M. H., Giangola, J. P., & Balogh, J. (2004). *Voice user interface design*. Addison-Wesley Professional.

- Cohen, P. A., Kulik, J. A., & Kulik, C.-L. C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American educational research journal*, 19(2), 237–248.
- Colgrove, J., Abiola, S., & Mello, M. M. (2010). Hpv vaccination mandates—lawmaking amid political and scientific controversy. *New England Journal of Medicine*, 363(8), 785–791.
- Constantine, N. A. & Jerman, P. (2007). Acceptance of human papillomavirus vaccination among californian parents of daughters: a representative statewide analysis. *Journal of Adolescent Health*, 40(2), 108–115.
- Core, M. G., Moore, J. D., & Zinn, C. (2003). The role of initiative in tutorial dialogue. In *Proceedings of the tenth conference on European chapter of the Association for Computational Linguistics-Volume 1* (pp. 67–74).: Association for Computational Linguistics.
- Craig, S. D., Driscoll, D. M., & Gholson, B. (2004). Constructing knowledge from dialog in an intelligent tutoring system: Interactive learning, vicarious learning, and pedagogical agents. *Journal of Educational Multimedia and Hypermedia*, 13(2), 163–183.
- Cunningham, R. M. & Boom, J. A. (2013). Telling Stories of Vaccine-Preventable Diseases: Why it Works. *South Dakota Medicine*, (pp.6).
- Dahlström, L. A., Tran, T. N., Lundholm, C., Young, C., Sundström, K., & Sparén, P. (2010). Attitudes to hpv vaccination among parents of children aged 12-15 years—a population-based survey in sweden. *International Journal of Cancer*, 126(2), 500–507.

- Daley, E. M., Marhefka, S. L., Buhi, E. R., Vamos, C. A., Hernandez, N. D., & Giuliano, A. R. (2010). Human papillomavirus vaccine intentions among men participating in a human papillomavirus natural history study versus a comparison sample. *Sexually transmitted diseases*, 37(10), 644.
- Daley, M. F., Liddon, N., Crane, L. A., Beaty, B. L., Barrow, J., Babbel, C., Markowitz, L. E., Dunne, E. F., Stokley, S., Dickinson, L. M., et al. (2006). A national survey of pediatrician knowledge and attitudes regarding human papillomavirus vaccination. *Pediatrics*, 118(6), 2280–2289.
- Damianos, L., Loehr, D., Burke, C., Hansen, S., & Vismeg, M. (2003). The msia experiment: Using speech to enhance human performance on a cognitive task. *International Journal of Speech Technology*, 6(2), 133–144.
- Damljanovic, D., Agatonovic, M., & Cunningham, H. (2011). Freya: An interactive way of querying linked data using natural language. In *The Semantic Web: ESWC 2011 Workshops* (pp. 125–138).: Springer.
- Davidoff, F. (1997). Time. *Annals of Internal Medicine*, 127(6), 483–485.
- De La Paz, S. & Graham, S. (1997). Effects of dictation and advanced planning instruction on the composing of students with writing and learning problems. *Journal of Educational Psychology*, 89(2), 203.

- Del Corro, L. & Gemulla, R. (2013). ClausIE: clause-based open information extraction. In *Proceedings of the 22nd international conference on World Wide Web* (pp. 355–366).: International World Wide Web Conferences Steering Committee.
- Delichatsios, H. K., Friedman, R. H., Glanz, K., Tennstedt, S., Smigelski, C., Pinto, B. M., Kelley, H., & Gillman, M. W. (2001). Randomized trial of a “talking computer” to improve adults’ eating habits. *American Journal of Health Promotion*, 15(4), 215–224.
- Dempsey, A. F., Abraham, L. M., Dalton, V., & Ruffin, M. (2009). Understanding the reasons why mothers do or do not have their adolescent daughters vaccinated against human papillomavirus. *Annals of Epidemiology*, 19(8), 531 – 538.
- Dempsey, A. F., Brewer, S. E., Pyrzanowski, J., Sevic, C., & O’leary, S. T. (2015). Acceptability of human papillomavirus vaccines among women older than 26 years. *Vaccine*, 33(13), 1556–1561.
- Dempsey, A. F., Fuhrel-Forbis, A., & Konrath, S. (2014). Use of the Carolina HPV Immunization Attitudes and Beliefs Scale (CHIAS) in Young Adult Women. *PLoS ONE*, 9(6), e100193.
- Dempsey, A. F., Zimet, G. D., Davis, R. L., & Koutsky, L. (2006). Factors that are associated with parental acceptance of human papillomavirus vaccines: a randomized intervention study of written information about hpv. *Pediatrics*, 117(5), 1486–1493.

- Deschler, D. G., Richmon, J. D., Khariwala, S. S., Ferris, R. L., & Wang, M. B. (2014). The “new” head and neck cancer patient—young, nonsmoker, nondrinker, and hpv positive evaluation. *Otolaryngology–Head and Neck Surgery*, 151(3), 375–380.
- Dinh, T. A., Rosenthal, S. L., Doan, E. D., Trang, T., Pham, V. H., Tran, B. D., Tran, V. D., Phan, G. B., Chu, H., & Breitkopf, C. R. (2007). Attitudes of mothers in da nang, vietnam toward a human papillomavirus vaccine. *Journal of Adolescent Health*, 40(6), 559–563.
- Dividino, R. Q., Romanelli, M., Sonntag, D., et al. (2008). Semiotic-based ontology evaluation tool (s-ontoeval). In *LREC*.
- D’Mello, S. K., Graesser, A., & King, B. (2010). Toward spoken human–computer tutorial dialogues. *Human–Computer Interaction*, 25(4), 289–323.
- Donadiki, E., Jiménez-García, R., Hernández-Barrera, V., Sourtzi, P., Carrasco-Garrido, P., de Andrés, A. L., Jimenez-Trujillo, I., & Velonakis, E. (2014). Health belief model applied to non-compliance with hpv vaccine among female university students. *Public health*, 128(3), 268–273.
- Donadiki, E. M., Jiménez-García, R., Hernández-Barrera, V., Carrasco-Garrido, P., de Andrés, A. L., & Velonakis, E. G. (2012). Human papillomavirus vaccination coverage among greek higher education female students and predictors of vaccine uptake. *Vaccine*, 30(49), 6967–6970.

- Dorell, C., Yankey, D., & Strasser, S. (2011a). Parent-reported reasons for nonreceipt of recommended adolescent vaccinations, national immunization survey—teen, 2009. *Clinical Pediatrics*, 50(12), 1116–1124.
- Dorell, C. G., Yankey, D., Santibanez, T. A., & Markowitz, L. E. (2011b). Human papillomavirus vaccination series initiation and completion, 2008–2009. *Pediatrics*, 128(5), 830–839.
- Downs, L. S., Smith, J. S., Scarinci, I., Flowers, L., & Parham, G. (2008). The disparity of cervical cancer in diverse populations. *Gynecologic oncology*, 109(2), S22–S30.
- Du, P., Camacho, F., McCall-Hosenfeld, J., Lengerich, E., Meyers, C. M., & Christensen, N. D. (2015). Human papillomavirus vaccination among adults and children in 5 us states. *Journal of Public Health Management and Practice*, 21(6), 573–583.
- Dubé, E., Gagnon, D., Ouakki, M., Bettinger, J. A., Guay, M., Halperin, S., Wilson, K., Graham, J., Witteman, H. O., MacDonald, S., et al. (2016). Understanding vaccine hesitancy in canada: Results of a consultation study by the canadian immunization research network. *PloS one*, 11(6), e0156118.
- Dunn, A. G., Leask, J., Zhou, X., Mandl, K. D., & Coiera, E. (2015). Associations between exposure to and expression of negative opinions about human papillomavirus vaccines on social media: an observational study. *Journal of medical Internet research*, 17(6).

- Dunne, E. F., Nielson, C. M., Stone, K. M., Markowitz, L. E., & Giuliano, A. R. (2006). Prevalence of hpv infection among men: a systematic review of the literature. *Journal of Infectious Diseases*, 194(8), 1044–1057.
- Dunne, E. F., Unger, E. R., Sternberg, M., McQuillan, G., Swan, D. C., Patel, S. S., & Markowitz, L. E. (2007). Prevalence of hpv infection among females in the united states. *Jama*, 297(8), 813–819.
- Eclipse Foundation (2019). Eclipse RDF4J. <https://rdf4j.eclipse.org/>.
- Ekman, P., Friesen, W. V., & Ellsworth, P. (2013). *Emotion in the human face: Guidelines for research and an integration of findings*. Elsevier.
- Esposito, S., Bosis, S., Pelucchi, C., Begliatti, E., Rognoni, A., Bellasio, M., Tel, F., Consolo, S., & Principi, N. (2007). Pediatrician knowledge and attitudes regarding human papillomavirus disease and its prevention. *Vaccine*, 25(35), 6437–6446.
- Etter, D. J., Zimet, G. D., & Rickert, V. I. (2012). Human papillomavirus vaccine in adolescent women: a 2012 update. *Current Opinion in Obstetrics and Gynecology*, 24(5), 305–310.
- Evans, G. & Bostrom, A. (2002). The evolution of vaccine risk communication in the united states: 1982–2002. *National Institute of Allergy and Infectious Diseases, ed. Jordan Report 20th anniversary: Accelerated Development of Vaccines*, (pp. 58–71).
- Evens, M. & Michael, J. (2006). One-on-one tutoring by humans and machines. *Computer Science Department, Illinois Institute of Technology*.

- Fabry, P., Gagneur, A., & Pasquier, J.-C. (2011). Determinants of a (h1n1) vaccination: cross-sectional study in a population of pregnant women in quebec. *Vaccine*, 29(9), 1824–1829.
- Fader, A., Soderland, S., & Etzioni, O. (2011). Identifying relations for open information extraction. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (pp. 1535–1545).: Association for Computational Linguistics.
- Fang, C. Y., Coups, E. J., & Heckman, C. J. (2010). Behavioral correlates of hpv vaccine acceptability in the 2007 health information national trends survey (hints). *Cancer Epidemiology Biomarkers & Prevention*, 19(2), 319–326.
- Farzanfar, R., Finkelstein, J., & Friedman, R. H. (2004). Testing the usability of two automated home-based patient-management systems. *Journal of medical systems*, 28(2), 143–153.
- Feemster, K. A., Winters, S. E., Fiks, A. G., Kinsman, S., & Kahn, J. A. (2008). Pediatricians' intention to recommend human papillomavirus (hpv) vaccines to 11-to 12-year-old girls postlicensing. *Journal of Adolescent Health*, 43(4), 408–411.
- Ferris, D. G., Waller, J. L., Miller, J., Patel, P., Jackson, L., Price, G. A., & Wilson, C. (2008). Men's attitudes toward receiving the human papillomavirus vaccine. *Journal of Lower Genital Tract Disease*, 12(4), 276–281.
- Ferris, D. G., Waller, J. L., Miller, J., Patel, P., Price, G. A., Jackson, L., & Wilson, C. (2009). Variables associated with human papillomavirus (hpv) vaccine acceptance by men. *The Journal of the American Board of Family Medicine*, 22(1), 34–42.



- Fiks, A. G., Grundmeier, R. W., Mayne, S., Song, L., Feemster, K., Karavite, D., Hughes, C. C., Massey, J., Keren, R., Bell, L. M., et al. (2013). Effectiveness of decision support for families, clinicians, or both on hpv vaccine receipt. *Pediatrics*, 131(6), 1114–1124.
- Finlayson, M. (2014). Java libraries for accessing the princeton wordnet: Comparison and evaluation. In *Proceedings of the Seventh Global Wordnet Conference* (pp. 78–85).
- Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*, 10(4), 507–521.
- Fisher, W. A. (2012). Understanding human papillomavirus vaccine uptake. *Vaccine*, 30, F149–F156.
- Forster, A. S., Marlow, L. A., Wardle, J., Stephenson, J., & Waller, J. (2010). Understanding adolescents' intentions to have the hpv vaccine. *Vaccine*, 28(7), 1673–1676.
- Fraser, N. M. & Gilbert, G. N. (1991). Simulating speech systems. *Computer Speech & Language*, 5(1), 81–99.
- Freed, G. L., Clark, S. J., Butchart, A. T., Singer, D. C., & Davis, M. M. (2010). Parental vaccine safety concerns in 2009. *Pediatrics*, 125(4), 654–659.
- Friedman, R. H., Kazis, L. E., Jette, A., Smith, M. B., Stollerman, J., Torgerson, J., & Carey, K. (1996). A telecommunications system for monitoring and counseling patients with hypertension: impact on medication adherence and blood pressure control. *American journal of hypertension*, 9(4), 285–292.

- Fu, L. Y., Bonhomme, L.-A., Cooper, S. C., Joseph, J. G., & Zimet, G. D. (2014). Educational interventions to increase {HPV} vaccination acceptance: A systematic review. *Vaccine*, 32(17), 1901–1920.
- Funk, A., Tablan, V., Bontcheva, K., Cunningham, H., Davis, B., & Handschuh, S. (2007). Clone: Controlled language for ontology editing. In *The Semantic Web* (pp. 142–155). Springer.
- Gallan, A. S., Jarvis, C. B., Brown, S. W., & Bitner, M. J. (2013). Customer positivity and participation in services: an empirical test in a health care context. *Journal of the Academy of Marketing Science*, 41(3), 338–356.
- Gangemi, A., Catenacci, C., Ciaramita, M., & Lehmann, J. (2006). Qood grid: A metaontology-based framework for ontology evaluation and selection. In *Proceedings of the EON'2006 Workshop*.
- Garland, S., Hernandez-Avila, M., Wheeler, C., Perez, G., Harper, D., Leodolter, S., Tang, G., Ferris, D., Steben, M., Bryan, J., et al. (2007). Females united to unilaterally reduce endo/ectocervical disease (future) i investigators quadrivalent vaccine against human papillomavirus to prevent anogenital diseases. *N Engl J Med*, 356(19), 1928–1943.
- Garland, S. & Smith, J. (2010). Human papillomavirus vaccines. *Drugs*, 70(9), 1079–1098.
- Gashteovski, K., Gemulla, R., & Del Corro, L. (2017). Minie: minimizing facts in open information extraction. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing* (pp. 2630–2640).

- Gatt, A. & Reiter, E. (2009). Simplenlg: A realisation engine for practical applications. In *Proceedings of the 12th European Workshop on Natural Language Generation (ENLG 2009)* (pp. 90–93).
- Gelman, A., Miller, E., Schwarz, E. B., Akers, A. Y., Jeong, K., & Borrero, S. (2013). Racial disparities in human papillomavirus vaccination: does access matter? *Journal of Adolescent Health*, 53(6), 756–762.
- Gerend, M. A. & Barley, J. (2009). Human papillomavirus vaccine acceptability among young adult men. *Sexually transmitted diseases*, 36(1), 58–62.
- Gerend, M. A. & Shepherd, J. E. (2012). Predicting human papillomavirus vaccine uptake in young adult women: Comparing the health belief model and theory of planned behavior. *Annals of Behavioral Medicine*, 44(2), 171–180.
- Giede, C., McFadden, L., Komonoski, P., Agrawal, A., Stauffer, A., & Pierson, R. (2010). Acceptability of hpv vaccination among women attending the university of saskatchewan student health services. *Gynecologic Oncology*, 116(3), 598.
- Ginzburg, J. & Fernández, R. (2010). Computational models of dialogue. *Handbook of Computational Linguistics and Natural Language*, Oxford. Blackwell.
- Girardi, R. (2010). Guiding ontology learning and population by knowledge system goals. In *KEOD* (pp. 480–484).

Giuliano, A. R., Lee, J.-H., Fulp, W., Villa, L. L., Lazcano, E., Papenfuss, M. R., Abrahamsen, M., Salmeron, J., Anic, G. M., Rollison, D. E., et al. (2011). Incidence and clearance of genital human papillomavirus infection in men (him): a cohort study. *The Lancet*, 377(9769), 932–940.

Glanz, K., Shigaki, D., Farzanfar, R., Pinto, B., Kaplan, B., & Friedman, R. H. (2003). Participant reactions to a computerized telephone system for nutrition and exercise counseling. *Patient education and counseling*, 49(2), 157–163.

Glimm, B., Horrocks, I., Motik, B., Stoilos, G., & Wang, Z. (2014). Hermit: an owl 2 reasoner. *Journal of Automated Reasoning*, 53(3), 245–269.

Goff, S. L., Mazor, K. M., Gagne, S. J., Corey, K. C., & Blake, D. R. (2011). Vaccine counseling: A content analysis of patient–physician discussions regarding human papilloma virus vaccine. *Vaccine*, 29(43), 7343–7349.

Goldie, S. J., O’Shea, M., Diaz, M., & Kim, S.-Y. (2008). Benefits, cost requirements and cost-effectiveness of the hpv16, 18 vaccine for cervical cancer prevention in developing countries: policy implications. *Reproductive health matters*, 16(32), 86–96.

Gómez-Pérez, A. (2004). Ontology evaluation,[in:] staab s., studer r.(eds.): Handbook on ontologies.

- Gomez-Perez, A., Fernández-López, M., & Corcho, O. (2006). *Ontological Engineering: with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web*. Springer Science & Business Media.
- Gorenstein, C. (2015). Bill maher's bizarre anti-vaccine rant: Stop calling these people "kooks and liars". *Salon*.
- Gostin, L. O. & DeAngelis, C. D. (2007). Mandatory hpv vaccination: public health vs private wealth. *Jama*, 297(17), 1921–1923.
- Gottlieb, S. L., Brewer, N. T., Sternberg, M. R., Smith, J. S., Ziarnowski, K., Liddon, N., & Markowitz, L. E. (2009). Human papillomavirus vaccine initiation in an area with elevated rates of cervical cancer. *Journal of Adolescent Health*, 45(5), 430–437.
- Gottlieb, S. L., Low, N., Newman, L. M., Bolan, G., Kamb, M., & Broutet, N. (2014). Toward global prevention of sexually transmitted infections (stis): the need for sti vaccines. *Vaccine*, 32(14), 1527–1535.
- Gowda, C., Carlos, R. C., Butchart, A. T., Singer, D. C., Davis, M. M., Clark, S. J., & Dempsey, A. F. (2012). Chias: a standardized measure of parental hpv immunization attitudes and beliefs and its associations with vaccine uptake. *Sexually transmitted diseases*, 39(6), 475–481.

- Grabel, M., Reutzel, T. J., Wang, S., Rubin, R., Leung, V., Ordonez, A., Wong, M., & Jordan, E. (2013). Hpv and hpv vaccines: the knowledge levels, opinions, and behavior of parents. *Journal of community health*, 38(6), 1015–1021.
- Graesser, A. C. & Olde, B. A. (2003). How does one know whether a person understands a device? the quality of the questions the person asks when the device breaks down. *Journal of Educational Psychology*, 95(3), 524.
- Green Jr, B. F., Wolf, A. K., Chomsky, C., & Laughery, K. (1961). Baseball: an automatic question-answerer. In *Papers presented at the May 9-11, 1961, western joint IRE-AIEE-ACM computer conference* (pp. 219–224).: ACM.
- Greenfield, K. T. (2010). The autism debate: Who’s afraid of jenny mccarthy? *Time Magazine*.
- Groth, P., Gibson, A., & Velterop, J. (2010). The anatomy of a nanopublication. *Information Services & Use*, 30(1-2), 51–56.
- Groth, P., Schultes, E., Thompson, M., Tatum, Z., & Dumontier, M. (2013). Nanopublication Guidelines. Concept Web Alliance Working Draft.
- Group, F. I. S. et al. (2010). Four year efficacy of prophylactic human papillomavirus quadrivalent vaccine against low grade cervical, vulvar, and vaginal intraepithelial neoplasia and anogenital warts: randomised controlled trial. *BMJ: British Medical Journal*, 341.
- Groza, A., Dragoste, I., Sincai, I., Jimborean, I., & Moraru, V. (2014). An ontology selection and ranking system based on the analytic hierarchy process. In *Symbolic and Numeric*

- Algorithms for Scientific Computing (SYNASC), 2014 16th International Symposium on* (pp. 293–300).: IEEE.
- Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing? *International Journal of Human-Computer Studies*, 43(5–6), 907–928.
- Guadagno, R. E., Blascovich, J., Bailenson, J. N., & Mccall, C. (2007). Virtual humans and persuasion: The effects of agency and behavioral realism. *Media Psychology*, 10(1), 1–22.
- Hadzic, M., Chang, E., Dillon, T., & Wongthongtham, P. (2009). *Ontology-based multi-agent systems*. Springer.
- Hanson, K. E., McLean, H. Q., Belongia, E. A., Stokley, S., McNeil, M. M., Gee, J., & VanWormer, J. J. (2019). Sociodemographic and clinical correlates of human papillomavirus vaccine attitudes and receipt among wisconsin adolescents. *Papillomavirus Research*, 8, 100168.
- Hariri, S., Unger, E. R., Sternberg, M., Dunne, E. F., Swan, D., Patel, S., & Markowitz, L. E. (2011). Prevalence of genital human papillomavirus among females in the united states, the national health and nutrition examination survey, 2003–2006. *Journal of Infectious Diseases*, 204(4), 566–573.
- Harris, S. & Biermann, A. W. (2002). Mouse selection versus voice selection of menu items. *International Journal of Speech Technology*, 5(4), 389–402.
- Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). Emotional contagion. *Current directions in psychological science*, 2(3), 96–100.

- Hemminki, K., Dong, C., & Frisch, M. (2000). Tonsillar and other upper aerodigestive tract cancers among cervical cancer patients and their husbands. *European Journal of Cancer Prevention*, 9(6), 433–437.
- Henderson, L., Clements, A., Damery, S., Wilkinson, C., Austoker, J., & Wilson, S. (2011). ‘a false sense of security’? understanding the role of the hpv vaccine on future cervical screening behaviour: a qualitative study of uk parents and girls of vaccination age. *Journal of medical screening*, 18(1), 41–45.
- Hendry, M., Lewis, R., Clements, A., Damery, S., & Wilkinson, C. (2013). “hpv? never heard of it!”: A systematic review of girls’ and parents’ information needs, views and preferences about human papillomavirus vaccination. *Vaccine*, 31(45), 5152–5167.
- Hernandez, B. Y., Wilkens, L., Thompson, P., Shvetso, Y., Goodman, M., Ning, L., & Kaopua, L. (2010). Acceptability of prophylactic human papillomavirus vaccination among adult men. *Human vaccines*, 6(6), 467–475.
- Herzog, T. J., Huh, W. K., & Einstein, M. H. (2010). How does public policy impact cervical screening and vaccination strategies? *Gynecologic oncology*, 119(2), 175–180.
- Hilton, S. & Smith, E. (2011). “i thought cancer was one of those random things. i didn’t know cancer could be caught...”: Adolescent girls’ understandings and experiences of the hpv programme in the uk. *Vaccine*, 29(26), 4409–4415.



- Hongpaisanwiwat, C. & Lewis, M. (2003). Attentional effect of animated character. *Proceedings of the human-computer interaction. iOS Press*, (pp. 423–30).
- Hopfer, S. & Clippard, J. R. (2010). College women's hpv vaccine decision narratives. *Qualitative Health Research*.
- Hopkins, T. G. & Wood, N. (2013). Female human papillomavirus (hpv) vaccination: global uptake and the impact of attitudes. *Vaccine*, 31(13), 1673–1679.
- Horridge, M. & Bechhofer, S. (2011). The owl api: A java api for owl ontologies. *Semantic Web*, 2(1), 11–21.
- Horrocks, I., Patel-Schneider, P. F., Boley, H., Tabet, S., Grosof, B., Dean, M., et al. (2004). Swrl: A semantic web rule language combining owl and ruleml. *W3C Member submission*, 21(79), 1–31.
- Hughes, J., Cates, J. R., Liddon, N., Smith, J. S., Gottlieb, S. L., & Brewer, N. T. (2009). Disparities in how parents are learning about the human papillomavirus vaccine. *Cancer Epidemiology Biomarkers & Prevention*, 18(2), 363–372.
- Humiston, S. G., Albertin, C., Schaffer, S., Rand, C., Shone, L. P., Stokley, S., & Szilagyi, P. G. (2009). Health care provider attitudes and practices regarding adolescent immunizations: a qualitative study. *Patient education and counseling*, 75(1), 121–127.
- InfoWars (2017). Vaccines exposed: The hidden crime against children.  
[www.infowars.com/vaccines-exposed-the-hidden-crime-against-children](http://www.infowars.com/vaccines-exposed-the-hidden-crime-against-children).

- Ishibashi, K. L., Koopmans, J., Curlin, F. A., Alexander, K. A., & Ross, L. F. (2008). Paediatricians' attitudes and practices towards hpv vaccination. *Acta Paediatrica*, 97(11), 1550–1556.
- Jain, N., Euler, G. L., Shefer, A., Lu, P., Yankey, D., & Markowitz, L. (2009). Human papillomavirus (hpv) awareness and vaccination initiation among women in the united states, national immunization survey—adult 2007. *Preventive medicine*, 48(5), 426–431.
- Janz, N. K. & Becker, M. H. (1984). The health belief model: A decade later. *Health Education & Behavior*, 11(1), 1–47.
- Jarvis, K. L., Friedman, R. H., Heeren, T., & Cullinane, P. M. (1997). Older women and physical activity: using the telephone to walk. *Women's Health Issues*, 7(1), 24–29.
- Javanbakht, M., Stahlman, S., Walker, S., Gottlieb, S., Markowitz, L., Liddon, N., Plant, A., & Guerry, S. (2012). Provider perceptions of barriers and facilitators of hpv vaccination in a high-risk community. *Vaccine*, 30(30), 4511–4516.
- Jemal, A., Simard, E. P., Dorell, C., Noone, A.-M., Markowitz, L. E., Kohler, B., Ehemann, C., Saraiya, M., Bandi, P., Saslow, D., et al. (2013). Annual report to the nation on the status of cancer, 1975–2009, featuring the burden and trends in human papillomavirus (hpv)–associated cancers and hpv vaccination coverage levels. *Journal of the National Cancer Institute*, (pp. djs491).

- Jeudin, P., Liveright, E., del Carmen, M. G., & Perkins, R. B. (2014). Race, ethnicity, and income factors impacting human papillomavirus vaccination rates. *Clinical therapeutics*, 36(1), 24–37.
- Jianliang, X. & Xiaowei, M. (2008). A web-based ontology evaluation system. In *Advanced Language Processing and Web Information Technology, 2008. ALPIT'08. International Conference on* (pp. 104–107).: IEEE.
- Jimboean, I. & Groza, A. (2014). Ranking ontologies in the ontology building competition boc 2014. In *Intelligent Computer Communication and Processing (ICCP), 2014 IEEE International Conference on* (pp. 75–82).: IEEE.
- Johnson, W. L., Rickel, J. W., & Lester, J. C. (2000). Animated pedagogical agents: Face-to-face interaction in interactive learning environments. *International Journal of Artificial intelligence in education*, 11(1), 47–78.
- Jokinen, K. & McTear, M. (2009). Spoken dialogue systems. *Synthesis Lectures on Human Language Technologies*, 2(1), 1–151.
- Jones, M. & Cook, R. (2008). Intent to receive an hpv vaccine among university men and women and implications for vaccine administration. *Journal of American College Health*, 57(1), 23–32.

- Juntasopeepun, P., Suwan, N., Phianmongkhol, Y., & Srisomboon, J. (2012). Factors influencing acceptance of human papillomavirus vaccine among young female college students in thailand. *International Journal of Gynecology & Obstetrics*, 118(3), 247–250.
- Jurafsky, D. & Martin, J. H. (2009). *Speech and Language Processing*. Pearson.
- Kabanza, F., Bisson, G., Charneau, A., & Jang, T.-S. (2006). Implementing tutoring strategies into a patient simulator for clinical reasoning learning. *Artificial Intelligence in Medicine*, 38(1), 79–96.
- Kaelbling, L. P. & Rosenschein, S. J. (1990). Action and planning in embedded agents. *Robotics and autonomous systems*, 6(1-2), 35–48.
- Kahn, J. A., Cooper, H. P., Vadaparampil, S. T., Pence, B. C., Weinberg, A. D., LoCoco, S. J., & Rosenthal, S. L. (2009). Human papillomavirus vaccine recommendations and agreement with mandated human papillomavirus vaccination for 11-to-12-year-old girls: a statewide survey of texas physicians. *Cancer Epidemiology Biomarkers & Prevention*, 18(8), 2325–2332.
- Kahn, J. A., Rosenthal, S. L., Tissot, A. M., Bernstein, D. I., Wetzel, C., & Zimet, G. D. (2007). Factors influencing pediatricians' intention to recommend human papillomavirus vaccines. *Ambulatory Pediatrics*, 7(5), 367–373.
- Kahn, J. A., Xu, J., Zimet, G. D., Liu, N., Gonin, R., Dillard, M. E., Squires, K., for HIV/AIDS Interventions, A. T. N., et al. (2012). Risk perceptions after human

- papillomavirus vaccination in hiv-infected adolescents and young adult women. *Journal of Adolescent Health*, 50(5), 464–470.
- Kahn, J. A., Zimet, G. D., Bernstein, D. I., Riedesel, J. M., Lan, D., Huang, B., & Rosenthal, S. L. (2005). Pediatricians' intention to administer human papillomavirus vaccine: the role of practice characteristics, knowledge, and attitudes. *Journal of Adolescent Health*, 37(6), 502–510.
- Kamimura, A., Trinh, H. N., Weaver, S., Chernenko, A., Wright, L., Stoddard, M., Nourian, M. M., & Nguyen, H. (2018). Knowledge and beliefs about hpv among college students in vietnam and the united states. *Journal of infection and public health*, 11(1), 120–125.
- Kang, H. S. & Moneyham, L. (2010). Attitudes toward and intention to receive the human papilloma virus (hpv) vaccination and intention to use condoms among female korean college students. *Vaccine*, 28(3), 811–816.
- Kata, A. (2010). A postmodern pandora's box: anti-vaccination misinformation on the internet. *Vaccine*, 28(7), 1709–1716.
- Kata, A. (2012). Anti-vaccine activists, web 2.0, and the postmodern paradigm—an overview of tactics and tropes used online by the anti-vaccination movement. *Vaccine*, 30(25), 3778–3789.

- Katz, M. L., Krieger, J. L., & Roberto, A. J. (2011). Human papillomavirus (hpv): college male's knowledge, perceived risk, sources of information, vaccine barriers and communication. *Journal of men's health*, 8(3), 175–184.
- Katz, M. L., Reiter, P. L., Heaner, S., Ruffin, M. T., Post, D. M., & Paskett, E. D. (2009). Acceptance of the hpv vaccine among women, parents, community leaders, and healthcare providers in ohio appalachia. *Vaccine*, 27(30), 3945–3952.
- Kaufmann, E. & Bernstein, A. (2007). *How useful are natural language interfaces to the semantic web for casual end-users?* Springer.
- Kaufmann, E., Bernstein, A., & Fischer, L. (2007). Nlp-reduce: A naive but domain-independent natural language interface for querying ontologies. In *Proceedings of the 4th European Semantic Web Conference (ESWC 2007)*: ESWC Zurich.
- Kaufmann, E., Bernstein, A., & Zumstein, R. (2006). Querix: A natural language interface to query ontologies based on clarification dialogs. In *5th International Semantic Web Conference (ISWC 2006)*(pp. 980–981): Citeseer.
- Kazama, J., Makino, T., Ohta, Y., & Tsujii, J. (2002). Tuning support vector machines for biomedical named entity recognition. In *Proceedings of the ACL-02 workshop on Natural language processing in the biomedical domain-Volume 3* (pp. 1–8): Association for Computational Linguistics.

- Kennedy, C. M., Powell, J., Payne, T. H., Ainsworth, J., Boyd, A., & Buchan, I. (2012). Active Assistance Technology for Health-Related Behavior Change: An Interdisciplinary Review. *Journal of Medical Internet Research*, 14(3).
- Kessels, S. J., Marshall, H. S., Watson, M., Braunack-Mayer, A. J., Reuzel, R., & Tooher, R. L. (2012). Factors associated with hpv vaccine uptake in teenage girls: a systematic review. *Vaccine*, 30(24), 3546–3556.
- Kester, L. M., Zimet, G. D., Fortenberry, J. D., Kahn, J. A., & Shew, M. L. (2013). A national study of hpv vaccination of adolescent girls: rates, predictors, and reasons for non-vaccination. *Maternal and child health journal*, 17(5), 879–885.
- Kim, J. J. & Goldie, S. J. (2009). Cost effectiveness analysis of including boys in a human papillomavirus vaccination programme in the united states. *Bmj*, 339.
- King, A., Friedman, R., Marcus, B., Napolitano, M., Castro, C., & Forsyth, L. (2003). Increasing regular physical activity via humans or automated technology: the chat trial. *Ann Behav Med*, 25, SIII.
- Knublauch, H., Hendler, J. A., & Idehen, K. (2011). SPIN - Overview and Motivation. <https://www.w3.org/Submission/spin-overview/>.
- Ko, E. M., Missmer, S., & Johnson, N. R. (2010). Physician attitudes and practice toward human papillomavirus vaccination. *Journal of lower genital tract disease*, 14(4), 339–345.

- Kobetz, E., Kornfeld, J., Vanderpool, R. C., Finney Rutten, L. J., Parekh, N., O'Bryan, G., & Menard, J. (2010a). Knowledge of hpv among united states hispanic women: opportunities and challenges for cancer prevention. *Journal of health communication*, 15(sup3), 22–29.
- Kobetz, E., Mendoza, A. D., Menard, J., Rutten, L. F., Diem, J., Barton, B., Kornfeld, J., & McKenzie, N. (2010b). One size does not fit all: Differences in hpv knowledge between haitian and african american women. *Cancer Epidemiology Biomarkers & Prevention*, 19(2), 366–370.
- Konstantinova, N. & Orasan, C. (2012). Interactive Question Answering. *Emerging Applications of Natural Language Processing: Concepts and New Research*, (pp. 149–169).
- Kontos, E. Z., Emmons, K. M., Puleo, E., & Viswanath, K. (2012). Contribution of communication inequalities to disparities in human papillomavirus vaccine awareness and knowledge. *American journal of public health*, 102(10), 1911–1920.
- Korstanje, M. (2019). SimMetrics. <https://github.com/Simmetrics/simmetrics>.
- Koshiol, J., Rutten, L. F., Moser, R. P., & Hesse, N. (2009). Knowledge of human papillomavirus: differences by self-reported treatment for genital warts and sociodemographic characteristics. *Journal of health communication*, 14(4), 331–345.
- Krakov, M. M., Jensen, J. D., Carcioppolo, N., Weaver, J., Liu, M., & Guntzville, L. M. (2015). Psychosocial predictors of human papillomavirus vaccination intentions for young women



18 to 26: Religiosity, morality, promiscuity, and cancer worry. *Women's Health Issues*, 25(2),

105 – 111.

Kravitz, L. & Ross, C. (1993). On are you gonna go my way [recorded by lenny  
kravitz]. In *On Are you gonna go my way [Compact Disc]*. Hoboken, New Jersey:  
Waterfront Studios.

Krawczyk, A., Stephenson, E., Perez, S., Lau, E., & Rosberger, Z. (2013). Deconstructing human  
papillomavirus (hpv) knowledge: Objective and perceived knowledge in males' intentions  
to receive the hpv vaccine. *American Journal of Health Education*, 44(1), 26–31.

Krupp, K., Marlow, L. A., Kielmann, K., Doddaiiah, N., Mysore, S., Reingold, A. L., &  
Madhivanan, P. (2010). Factors associated with intention-to-recommend human  
papillomavirus vaccination among physicians in mysore, india. *Journal of Adolescent  
Health*, 46(4), 379–384.

Kuhn, T., Chichester, C., Krauthammer, M., & Dumontier, M. (2015). Publishing without  
publishers: A decentralized approach to dissemination, retrieval, and archiving of data.  
In *International Semantic Web Conference* (pp. 656–672). Cham: Springer.

Kwan, T. T., Chan, K. K., Yip, A. M., Tam, K., Cheung, A. N., Young, P. M., Lee, P. W., & Ngan,  
H. Y. (2008). Barriers and facilitators to human papillomavirus vaccination among chinese  
adolescent girls in hong kong: a qualitative–quantitative study. *Sexually transmitted  
infections*, 84(3), 227–232.

- Lacey, C. J., Lowndes, C. M., & Shah, K. V. (2006). Burden and management of non-cancerous hpv-related conditions: Hpv-6/11 disease. *Vaccine*, 24, S35–S41.
- Lambert, J. (2019). Is measles here to stay? *NPR.org*.
- Lane, H. C. & VanLehn, K. (2005). Teaching the tacit knowledge of programming to novices with natural language tutoring. *Computer Science Education*, 15(3), 183–201.
- Lau, M., Lin, H., & Flores, G. (2012). Factors associated with human papillomavirus vaccine-series initiation and healthcare provider recommendation in us adolescent females: 2007 national survey of children's health. *Vaccine*, 30(20), 3112–3118.
- Laz, T. H., Rahman, M., & Berenson, A. B. (2012). An update on human papillomavirus vaccine uptake among 11–17 year old girls in the united states: National health interview survey, 2010. *Vaccine*, 30(24), 3534–3540.
- Laz, T. H., Rahman, M., & Berenson, A. B. (2013). Human papillomavirus vaccine uptake among 18-to 26-year-old women in the united states. *Cancer*, 119(7), 1386–1392.
- Leask, J., Jackson, C., Trevena, L., McCaffery, K., & Brotherton, J. (2009). Implementation of the australian hpv vaccination program for adult women: qualitative key informant interviews. *Vaccine*, 27(40), 5505–5512.
- Leddy, M. A., Anderson, B. L., Gall, S., & Schulkin, J. (2009). Obstetrician-gynecologists and the hpv vaccine: practice patterns, beliefs, and knowledge. *Journal of pediatric and adolescent gynecology*, 22(4), 239–246.

- Lenahan, J., Leonard, K., Nandra, S., Isaacs, C., Mathew, A., & Fisher, W. (2008). Women's knowledge, attitudes, and intentions concerning human papillomavirus vaccination: findings of a waiting room survey of obstetrics-gynaecology outpatients. *Journal of obstetrics and gynaecology Canada: JOGC= Journal d'obstetrique et gynecologie du Canada: JOGC*, 30(6), 489–499.
- Lenselink, C., Schmeink, C., Melchers, W., Massuger, L., Hendriks, J., Van Hamont, D., & Bekkers, R. (2008a). Young adults and acceptance of the human papillomavirus vaccine. *Public health*, 122(12), 1295–1301.
- Lenselink, C. H., Gerrits, M. M., Melchers, W. J., Massuger, L. F., van Hamont, D., & Bekkers, R. L. (2008b). Parental acceptance of human papillomavirus vaccines. *European Journal of Obstetrics & Gynecology and reproductive biology*, 137(1), 103–107.
- Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making. *Annual review of psychology*, 66, 799–823.
- Lewis, J. R. (2016). Standardized Questionnaires for Voice Interaction Design. *Voice Interaction Design*, 1(1), 16.
- Lewis, J. R. & Hardzinski, M. L. (2015). Investigating the psychometric properties of the Speech User Interface Service Quality questionnaire. *International Journal of Speech Technology*, 18(3), 479–487.

- Lewis, J. R. & Sauro, J. (2017). Can i leave this one out?: the effect of dropping an item from the sus. *Journal of Usability Studies*, 13(1), 38–46.
- Liddon, N., Hood, J., Wynn, B. A., & Markowitz, L. E. (2010). Acceptability of human papillomavirus vaccine for males: a review of the literature. *Journal of Adolescent Health*, 46(2), 113–123.
- Liddon, N. C., Hood, J. E., & Leichliter, J. S. (2012a). Intent to receive hpv vaccine and reasons for not vaccinating among unvaccinated adolescent and young women: findings from the 2006–2008 national survey of family growth. *Vaccine*, 30(16), 2676–2682.
- Liddon, N. C., Leichliter, J. S., & Markowitz, L. E. (2012b). Human papillomavirus vaccine and sexual behavior among adolescent and young women. *American journal of preventive medicine*, 42(1), 44–52.
- Liddy, E. D. (2001). Natural language processing. In *Encyclopedia of Library and Information Science*. Marcel Decker, Inc.
- Lin, R., Amith, M. P., Liang, C., Duan, R., Chen, Y., & Tao, C. (2018). Visualized Emotion Ontology: a model for representing visual cues of emotions. *BMC Medical Informatics and Decision Making*, 18(2), 64.
- Lin, Y., Harris, M. R., Manion, F. J., Eisenhauer, E., Zhao, B., Shi, W., Karnovsky, A., & He, Y. (2014). Development of a BFO-based informed consent ontology (ICO). *Bioinformatics*.

- Litman, D. J., Rosé, C. P., Forbes-Riley, K., VanLehn, K., Bhembé, D., & Silliman, S. (2006). Spoken versus typed human and computer dialogue tutoring. *IJ Artificial Intelligence in Education*, 16(2), 145–170.
- Liu, D., Yang, Y., Chen, Y., Zhu, H., Bayley, I., & Aldea, A. (2016). Evaluating the ontological semantic description of web services generated from algebraic specifications. In *Service-Oriented System Engineering (SOSE), 2016 IEEE Symposium on* (pp. 211–220).: IEEE.
- Lopez, V., Pasin, M., & Motta, E. (2005). Aqualog: An ontology-portable question answering system for the semantic web. In *The Semantic Web: Research and Applications* (pp. 546–562). Springer.
- Lopez, V., Uren, V., Sabou, M., & Motta, E. (2011). Is question answering fit for the semantic web?: a survey. *Semantic Web*, 2(2), 125–155.
- Lövheim, H. (2012). A new three-dimensional model for emotions and monoamine neurotransmitters. *Medical hypotheses*, 78(2), 341–348.
- Luperfoy, S. (2004). Retrofitting Synthetic Dialog Agents to Game Characters for Lifestyle Risk Training. In *AAAI Fall Symposium on Dialogue Systems for Health Communication*. Washington, DC.
- Lustria, M., Brown, L., & Davis, R. (2007). 10 years of consumer health informatics: What have we learned about how to design successful web-based interventions? In *National Communication Association* (pp.1).: National Communication Association.

- Madden, K., Nan, X., Briones, R., & Waks, L. (2012). Sorting through search results: a content analysis of hpv vaccine information online. *Vaccine*, 30(25), 3741–3746.
- Madhivanan, P., Krupp, K., Yashodha, M., Marlow, L., Klausner, J. D., & Reingold, A. L. (2009). Attitudes toward hpv vaccination among parents of adolescent girls in mysore, india. *Vaccine*, 27(38), 5203–5208.
- Maes, P. (1991). The agent network architecture (ana). *Acm sigart bulletin*, 2(4), 115–120.
- Manion, F. (2017). *A Life Cycle Approach to the Development and Validation of an Ontology of the U.S. Common Rule (45 C.F.R. §46)*. PhD thesis, The University of Texas Health Science Center at Houston.
- Manning, C., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S., & McClosky, D. (2014). The stanford corenlp natural language processing toolkit. In *Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations* (pp. 55–60).
- Markowitz, L. E., Dunne, E. F., Saraiya, M., for Disease Control, C., (CDC), P., et al. (2014). Human papillomavirus vaccination: recommendations of the advisory committee on immunization practices (acip). *MMWR Recomm Rep*, 63(RR-05), 1–30.
- Markowitz, L. E., Tsu, V., Deeks, S. L., Cubie, H., Wang, S. A., Vicari, A. S., & Brotherton, J. M. (2012). Human papillomavirus vaccine introduction—the first five years. *Vaccine*, 30, F139–F148.

- Marlow, L. (2011). Hpv vaccination among ethnic minorities in the uk: knowledge, acceptability and attitudes. *British journal of cancer*, 105(4), 486–492.
- Marlow, L. A., Forster, A. S., Wardle, J., & Waller, J. (2009a). Mothers' and adolescents' beliefs about risk compensation following hpv vaccination. *Journal of Adolescent Health*, 44(5), 446–451.
- Marlow, L. A., Waller, J., Evans, R. E., & Wardle, J. (2009b). Predictors of interest in hpv vaccination: A study of british adolescents. *Vaccine*, 27(18), 2483–2488.
- Marlow, L. A., Waller, J., & Wardle, J. (2007). Parental attitudes to pre-pubertal hpv vaccination. *Vaccine*, 25(11), 1945–1952.
- Marlow, L. A., Wardle, J., & Waller, J. (2009c). Attitudes to hpv vaccination among ethnic minority mothers in the uk: an exploratory qualitative study. *Human vaccines*, 5(2), 105–110.
- Martone, M. E., Gupta, A., & Ellisman, M. H. (2004). E-neuroscience: challenges and triumphs in integrating distributed data from molecules to brains. *Nature neuroscience*, 7(5), 467–472.
- Mays, R. M., Sturm, L. A., & Zimet, G. D. (2004). Parental perspectives on vaccinating children against sexually transmitted infections. *Social science & medicine*, 58(7), 1405–1413.
- McClelland, A. & Liamputtong, P. (2006). Knowledge and acceptance of human papillomavirus vaccination: perspectives of young australians living in melbourne, australia. *Sexual health*, 3(2), 95–101.

- McColl-Kennedy, J. R., Danaher, T. S., Gallan, A. S., Orsingher, C., Lervik-Olsen, L., & Verma, R. (2017). How do you feel today? managing patient emotions during health care experiences to enhance well-being. *Journal of Business Research*, 79, 247–259.
- McRae, A.-L., Brewer, N. T., Reiter, P. L., Gottlieb, S. L., & Smith, J. S. (2010a). The carolina hpv immunization attitudes and beliefs scale (chias): scale development and associations with intentions to vaccinate. *Sexually transmitted diseases*, 37(4), 234–239.
- McRae, A.-L., Reiter, P. L., & Brewer, N. T. (2012). Parents' internet use for information about hpv vaccine. *Vaccine*, 30(25), 3757–3762.
- McRae, A.-L., Reiter, P. L., Chantala, K., & Brewer, N. T. (2010b). Does framing human papillomavirus vaccine as preventing cancer in men increase vaccine acceptability? *Cancer Epidemiology Biomarkers & Prevention*, 19(8), 1937–1944.
- Michael, M. & Cheuvront, C. C. J. B. (1998). Health communication on the internet: an effective channel for health behavior change? *Journal of health communication*, 3(1), 71–79.
- Migneault, J., Rakowski, W., Friedman, R., Marks, L., & Clark, M. (2005). User engagement with an automated telephone intervention to improve on-time repeat mammography. *Annals of Behavioral Medicine*, D-33, S085.
- Migneault, J. P., Farzanfar, R., Wright, J. A., & Friedman, R. H. (2006). How to write health dialog for a talking computer. *Journal of biomedical informatics*, 39(5), 468–481.



- Miller, W. R. & Rollnick, S. (1991). *Motivational interviewing: Preparing people to change addictive behavior*. Guilford Press.
- Minnen, G., Carroll, J., & Pearce, D. (2001). Applied morphological processing of english. *Natural Language Engineering*, 7(3), 207–223.
- Montanaro, D. (2017). Despite the facts, Trump once again embraces vaccine skeptics. *NPR.org*.
- Montgomery, K. & Smith-Glasgow, M. E. (2012). Human papillomavirus and cervical cancer knowledge, health beliefs, and preventive practices in 2 age cohorts: a comparison study. *Gender medicine*, 9(1), S55–S66.
- Morales-Campos, D. Y., Markham, C. M., Peskin, M. F., & Fernandez, M. E. (2013). Hispanic mothers' and high school girls' perceptions of cervical cancer, human papilloma virus, and the human papilloma virus vaccine. *Journal of Adolescent Health*, 52(5), S69–S75.
- Moreno, R., Mayer, R. E., Spires, H. A., & Lester, J. C. (2001). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and instruction*, 19(2), 177–213.
- Mortensen, G. L. (2010). Drivers and barriers to acceptance of human-papillomavirus vaccination among young women: a qualitative and quantitative study. *BMC Public Health*, 10(1), 68.
- Moundridou, M. & Virvou, M. (2002). Evaluating the persona effect of an interface agent in a tutoring system. *Journal of computer assisted learning*, 18(3), 253–261.

- Müller, J. P. (1996). The agent architecture interrap. *The Design of Intelligent Agents: A Layered Approach*, (pp. 45–123).
- Mullins, T. L. K., Griffioen, A. M., Glynn, S., Zimet, G. D., Rosenthal, S. L., Fortenberry, J. D., & Kahn, J. A. (2013). Human papillomavirus vaccine communication: perspectives of 11–12 year-old girls, mothers, and clinicians. *Vaccine*, 31(42), 4894–4901.
- Munoz, N., Bosch, F. X., Castellsagué, X., Díaz, M., de Sanjose, S., Hammouda, D., Shah, K. V., & Meijer, C. J. (2004). Against which human papillomavirus types shall we vaccinate and screen? the international perspective. *International Journal of Cancer*, 111(2), 278–285.
- Musen, M. A. et al. (2015). The protégé project: a look back and a look forward. *AI matters*, 1(4), 4.
- Myers, M. G. & Pineda, D. (2009). Misinformation about vaccines. *Vaccines for biodefense and emerging and neglected diseases*, (pp. 255–270).
- Narayanan, S., Ananthakrishnan, S., Belvin, R., Ettelaie, E., Gandhe, S., Ganjavi, S., Georgiou, P. G., Hein, C. M., Kadambe, S., Knight, K., & others (2004). The Transonics spoken dialogue translator: an aid for English-Persian doctor-patient interviews. In *Working Notes of the AAAI Fall symposium on Dialogue Systems for Health Communication* (pp. 97–103).
- National Cancer Institute (2012).       Hpv and cancer.  
<http://www.cancer.gov/cancertopics/factsheet/Risk/HPV>.

- Naylor, M. D., Brooten, D. A., Campbell, R. L., Maislin, G., McCauley, K. M., & Schwartz, J. S. (2004). Transitional care of older adults hospitalized with heart failure: a randomized, controlled trial. *Journal of the American Geriatrics Society*, 52(5), 675–684.
- Newmann, S. J. & Garner, E. O. (2005). Social inequities along the cervical cancer continuum: a structured review. *Cancer Causes & Control*, 16(1), 63–70.
- Niccolai, L. M., Mehta, N. R., & Hadler, J. L. (2011). Racial/ethnic and poverty disparities in human papillomavirus vaccination completion. *American journal of preventive medicine*, 41(4), 428–433.
- Nicol, C., Duska, L., Hornsby, P., Pannone, A., & Bernheim, R. (2016). HPV vaccination rates and attitudes: A cross-sectional survey of college men and women. *Gynecologic Oncology*, 141, 112.
- Noakes, K., Yarwood, J., & Salisbury, D. (2006). Parental response to the introduction of a vaccine against human papilloma virus. *Human Vaccines*, 2(6), 243–248.
- Obrst, L., Ceusters, W., Mani, I., Ray, S., & Smith, B. (2007). The evaluation of ontologies. In *Semantic Web* (pp. 139–158). Springer.
- Offit, P. A. & Moser, C. A. (2011). *Vaccines & your child: separating fact from fiction*. Columbia University Press.

- Ogilvie, G. S., Remple, V. P., Marra, F., McNeil, S. A., Naus, M., Pielak, K., Ehlen, T., Dobson, S., Patrick, D. M., & Money, D. (2008). Intention of parents to have male children vaccinated with the human papillomavirus vaccine. *Sexually transmitted infections*, 84(4), 318–323.
- Olshen, E., Woods, E. R., Austin, S. B., Luskin, M., & Bauchner, H. (2005). Parental acceptance of the human papillomavirus vaccine. *Journal of Adolescent Health*, 37(3), 248–251.
- Opel, D. J., Diekema, D. S., & Marcuse, E. K. (2011). Assuring research integrity in the wake of wakefield. *BMJ*, 342, d2.
- Opel, D. J., Heritage, J., Taylor, J. A., Mangione-Smith, R., Salas, H. S., DeVere, V., Zhou, C., & Robinson, J. D. (2013). The Architecture of Provider-Parent Vaccine Discussions at Health Supervision Visits. *Pediatrics*, 132(6), 1037–1046.
- Opel, D. J., Mangione-Smith, R., Robinson, J. D., Heritage, J., DeVere, V., Salas, H. S., Zhou, C., & Taylor, J. A. (2015). The influence of provider communication behaviors on parental vaccine acceptance and visit experience. *American journal of public health*, 105(10), 1998–2004.
- Opel, D. J., Robinson, J. D., Heritage, J., Korfiatis, C., Taylor, J. A., & Mangione-Smith, R. (2012). Characterizing providers' immunization communication practices during health supervision visits with vaccine-hesitant parents: A pilot study. *Vaccine*, 30(7), 1269–1275.
- Ortony, A., Clore, G. L., & Collins, A. (1990). *The cognitive structure of emotions*. Cambridge university press.

- O'Connor, M., Knublauch, H., Tu, S., Grosz, B., Dean, M., Grosso, W., & Musen, M. (2005). Supporting rule system interoperability on the semantic web with swrl. In *International semantic web conference* (pp. 974–986).: Springer.
- Palefsky, J. M. (2010). Human papillomavirus-related disease in men: not just a women's issue. *Journal of Adolescent Health*, 46(4), S12–S19.
- paomedia (2019). User female alt icon. <http://www.iconarchive.com/show/small-n-flat-icons-by-paomedia/user-female-alt-icon.html>. Public Domain. Accessed 23 Apr 2019.
- Papirus Development Team (2019). Gnome robots icon. <http://www.iconarchive.com/show/papirus-apps-icons-by-papirus-team/gnome-robots-icon.html>. GNU General Public License v3.0. Accessed 23 Apr 2019.
- Parkin, D. M. & Bray, F. (2006). The burden of hpv-related cancers. *Vaccine*, 24, S11–S25.
- Patel, D. A., Zochowski, M., Peterman, S., Dempsey, A. F., Ernst, S., & Dalton, V. K. (2012). Human papillomavirus vaccine intent and uptake among female college students. *Journal of American College Health*, 60(2), 151–161.
- Pearce, C. F., Wolfe, L. M., DePasquale, S. E., & Breen, J. M. (2009). Physician knowledge and opinions about sexually transmitted infections and the human papillomavirus vaccine: a community based survey. *Tennessee medicine: journal of the Tennessee Medical Association*, 102(3), 33–5.

- Perera, R. & Nand, P. (2017). Recent advances in natural language generation: A survey and classification of the empirical literature. *Computing and Informatics*, 36(1), 1–32.
- Perkins, R. B. & Clark, J. A. (2012). What affects human papillomavirus vaccination rates? a qualitative analysis of providers' perceptions. *Women's Health Issues*, 22(4), e379–e386.
- Person, N. K., Graesser, A., Bautista, L., Mathews, E., Group, T. R., et al. (2001). Evaluating student learning gains in two versions of autotutor. *Artificial intelligence in education: AI-ED in the wired and wireless future*, (pp. 286–293).
- Petosa, R. & Jackson, K. (1991). Using the health belief model to predict safer sex intentions among adolescents. *Health Education & Behavior*, 18(4), 463–476.
- Petrovic, K., Burney, S., & Fletcher, J. (2011). The relationship of knowledge, health value and health self-efficacy with men's intentions to receive the human papillomavirus (hpv) vaccine. *Journal of Health Psychology*, 16(8), 1198–1207.
- Pinto, B. M., Friedman, R., Marcus, B. H., Kelley, H., Tennstedt, S., & Gillman, M. W. (2002). Effects of a computer-based, telephone-counseling system on physical activity. *American journal of preventive medicine*, 23(2), 113–120.
- Plutchik, R. (2001). The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American scientist*, 89(4), 344–350.

- Podolsky, R., Cremer, M., Atrio, J., Hochman, T., & Arslan, A. A. (2009). Hpv vaccine acceptability by latino parents: a comparison of us and salvadoran populations. *Journal of pediatric and adolescent gynecology*, 22(4), 205–215.
- Polkosky, M. D. (2005). Toward a social-cognitive psychology of speech technology: Affective responses to speech-based e-service.
- Polkosky, M. D. (2008). Machines as mediators: The challenge of technology for interpersonal communication theory and research. In *Mediated interpersonal communication* (pp. 48–71). Routledge.
- Portet, F., Reiter, E., Gatt, A., Hunter, J., Sripada, S., Freer, Y., & Sykes, C. (2009). Automatic generation of textual summaries from neonatal intensive care data. *Artificial Intelligence*, 173(7-8), 789–816.
- Pourat, N. & Jones, J. M. (2012). Role of insurance, income, and affordability in human papillomavirus vaccination. *The American journal of managed care*, 18(6), 320–330.
- Price, R. & Shanks, G. (2016). A semiotic information quality framework: development and comparative analysis. In *Enacting Research Methods in Information Systems* (pp. 219–250). Cham.: Palgrave Macmillan.
- Priest, H. M., Knowlden, A. P., & Sharma, M. (2015). Social cognitive theory predictors of human papillomavirus vaccination intentions of college men at a southeastern university. *International Quarterly of Community Health Education*, (pp. 0272684X15583289).

- Prochaska, J. O. & DiClemente, C. C. (2005). The transtheoretical approach. *Handbook of psychotherapy integration*, 2, 147–171.
- Procter, R. (2009). Definition of health informatics.  
<https://www.nlm.nih.gov/hsrinfo/informatics.html>.
- Quinn, G. P., Murphy, D., Malo, T. L., Christie, J., & Vadaparampil, S. T. (2012). A national survey about human papillomavirus vaccination: What we didn't ask, but physicians wanted us to know. *Journal of pediatric and adolescent gynecology*, 25(4), 254–258.
- Quintero Johnson, J., Sionean, C., & Scott, A. M. (2011). Exploring the presentation of news information about the hpv vaccine: a content analysis of a representative sample of us newspaper articles. *Health communication*, 26(6), 491–501.
- Racktoo, S. & Coverdale, G. (2009). Hpv? never heard of it. students and the hpv vaccine. *British Journal of School Nursing*, 4(7), 328–34.
- Rahati, A. (2012). *Persuasive and adaptive tutorial dialogues for a medical diagnosis tutoring system*. PhD thesis, University of Sherbrooke.
- Rahman, M., Islam, M., & Berenson, A. B. (2015). Differences in hpv immunization levels among young adults in various regions of the united states. *Journal of community health*, 40(3), 404–408.



- Rahman, M., Laz, T. H., McGrath, C., & Berenson, A. B. (2014). Correlates of human papillomavirus vaccine series completion among young adult female initiators. *Human vaccines & immunotherapeutics*, 10(8), 2163–2167.
- Rambout, L., Tashkandi, M., Hopkins, L., & Tricco, A. C. (2014). Self-reported barriers and facilitators to preventive human papillomavirus vaccination among adolescent girls and young women: a systematic review. *Preventive medicine*, 58, 22–32.
- Ramelson, H. Z., Bassey, B., & Friedman, R. H. (2003). The use of computer telephony to provide interactive health information. In *AMIA Annual Symposium Proceedings*, volume 2003 (pp. 539): American Medical Informatics Association.
- Ratanasiripong, N. T. (2012). A review of human papillomavirus (hpv) infection and hpv vaccine-related attitudes and sexual behaviors among college-aged women in the united states. *Journal of American College Health*, 60(6), 461–470.
- Rector, A., Drummond, N., Horridge, M., Rogers, J., Knublauch, H., Stevens, R., Wang, H., & Wroe, C. (2004). Owl pizzas: Practical experience of teaching owl-dl: Common errors & common patterns. In *International Conference on Knowledge Engineering and Knowledge Management* (pp. 63–81): Springer.
- Reis, L., Eisner, M., Kosary, C., Hankey, B., Miller, B., Clegg, L., Mariotto, A., Feuer, E., & Edwards, B. (2005). Seer cancer statistics review, 1975–2002. *National Cancer Institute, Bethesda, MD*, 6.

- Reiter, E. & Dale, R. (2000). *Building natural language generation systems*. Cambridge university press.
- Reiter, P. L., Brewer, N. T., Gottlieb, S. L., McRee, A.-L., & Smith, J. S. (2009). Parents' health beliefs and hpv vaccination of their adolescent daughters. *Social science & medicine*, 69(3), 475–480.
- Reiter, P. L., Brewer, N. T., & Smith, J. S. (2010). Human papillomavirus knowledge and vaccine acceptability among a national sample of heterosexual men. *Sexually transmitted infections*, 86(3), 241–246.
- Reiter, P. L., McRee, A.-L., Kadis, J. A., & Brewer, N. T. (2011). Hpv vaccine and adolescent males. *Vaccine*, 29(34), 5595–5602.
- Rickenberg, R. & Reeves, B. (2000). The effects of animated characters on anxiety, task performance, and evaluations of user interfaces. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 49–56).: ACM.
- Riedesel, J. M., Rosenthal, S. L., Zimet, G. D., Bernstein, D. I., Huang, B., Lan, D., & Kahn, J. A. (2005). Attitudes about human papillomavirus vaccine among family physicians. *Journal of pediatric and adolescent gynecology*, 18(6), 391–398.
- Rimer, B., Harper, H., & Witte, O. (2014). Accelerating hpv vaccine uptake: urgency for action to prevent cancer; a report to the president of the united states from the president's cancer panel. *National Cancer Institute, Bethesda, MD*.

- Robbins, S. C. C., Bernard, D., McCaffery, K., Brotherton, J., Garland, S., & Skinner, S. R. (2010). "is cancer contagious?": Australian adolescent girls and their parents: Making the most of limited information about hpv and hpv vaccination. *Vaccine*, 28(19), 3398–3408.
- Rosenthal, S., Weiss, T. W., Zimet, G. D., Ma, L., Good, M., & Vichnin, M. (2011). Predictors of hpv vaccine uptake among women aged 19–26: importance of a physician's recommendation. *Vaccine*, 29(5), 890–895.
- Rosenthal, S. L., Rupp, R., Zimet, G. D., Meza, H. M., Loza, M. L., Short, M. B., & Succop, P. A. (2008). Uptake of hpv vaccine: demographics, sexual history and values, parenting style, and vaccine attitudes. *Journal of Adolescent Health*, 43(3), 239–245.
- Roter, D. L., Erby, L. H., Larson, S., & Ellington, L. (2007). Assessing oral literacy demand in genetic counseling dialogue: preliminary test of a conceptual framework. *Social science & medicine*, 65(7), 1442–1457.
- Rowley, J. (2007). The wisdom hierarchy: representations of the dikw hierarchy. *Journal of information science*, 33(2), 163–180.
- Rubin, D. L., Shah, N. H., & Noy, N. F. (2007). Biomedical ontologies: a functional perspective. *Briefings in Bioinformatics*, 9(1), 75–90.
- Russell, J. A. & Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *Journal of research in Personality*, 11(3), 273–294.

- Saraiya, M., Rosser, J. I., & Cooper, C. P. (2012). Cancers that us physicians believe the hpv vaccine prevents: findings from a physician survey, 2009. *Journal of Women's Health*, 21(2), 111–117.
- Sauro, J. & Lewis, J. R. (2016). *Quantifying the user experience: Practical statistics for user research*. Morgan Kaufmann.
- Schiller, J. T., Castellsagué, X., & Garland, S. M. (2012). A review of clinical trials of human papillomavirus prophylactic vaccines. *Vaccine*, 30, F123–F138.
- Schmitz, M., Bart, R., Soderland, S., Etzioni, O., et al. (2012). Open language learning for information extraction. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning* (pp. 523–534).: Association for Computational Linguistics.
- Schmitz, M. & Thomson, S. (2013). Morpha stemmer. <https://github.com/knowitall/morpha>.
- Schnatz, P. F., Humphrey, K., & O'Sullivan, D. M. (2010). Assessment of the perceptions and administration of the human papillomavirus vaccine. *Journal of lower genital tract disease*, 14(2), 103–107.
- School of Biomedical Informatics Ontology Research Group, The University of Texas Health Science Center of Houston (2019). The Time Event Ontology (TEO).
- Schwarz, E. & Kurfess, F. (2011). Introduction to Ontologies with Protege. Department of Computer Science and Software Engineering - Cal Poly, San Luis Obispo.

- Schwarz, T. F., Spaczynski, M., Schneider, A., Wysocki, J., Galaj, A., Schulze, K., Poncelet, S. M., Catteau, G., Thomas, F., & Descamps, D. (2011). Persistence of immune response to hpv-16/18 aso4-adjuvanted cervical cancer vaccine in women aged 15-55 years. *Human vaccines*, 7(9), 958–965.
- Searle, J. R. (1976). A classification of illocutionary acts. *Language in society*, 5(01), 1–23.
- Sears, R. W. (2011). *The vaccine book: making the right decision for your child*. Little, Brown.
- Shah, N. H., Rubin, D. L., Supekar, K. S., & Musen, M. A. (2006). Ontology-based annotation and query of tissue microarray data. In *AMIA*.
- Sherris, J., Friedman, A., Wittet, S., Davies, P., Steben, M., & Saraiya, M. (2006). Education, training, and communication for hpv vaccines. *Vaccine*, 24, S210–S218.
- Shlyapnikova, A. (2019). Brick wall icon. <http://www.iconarchive.com/show/toolbar-2-icons-by-shlyapnikova/brick-wall-icon.html>. CC Attribution 4.0. Accessed 23 Apr 2019.
- Siegel, T. (2017). Controversial anti-vaccine doc 'vaxxed' gets secret cannes screening. *The Hollywood Reporter*.
- Sirin, E., Parsia, B., Grau, B. C., Kalyanpur, A., & Katz, Y. (2007). Pellet: A practical owl-dl reasoner. *Web Semantics: science, services and agents on the World Wide Web*, 5(2), 51–53.
- Smith, B. (2003). Ontology: philosophical and computational. *The Blackwell Guide to the Philosophy of Computing and Information*. Blackwell, Oxford.

- Smith, B., Ashburner, M., Rosse, C., Bard, J., Bug, W., Ceusters, W., Goldberg, L. J., Eilbeck, K., Ireland, A., Mungall, C. J., & others (2007). The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. *Nature biotechnology*, 25(11), 1251–1255.
- Smith, B. & Grenon, P. (2019). Basic Formal Ontology (BFO) | Home. <http://basic-formal-ontology.org/>.
- Songthap, A., Pitisuttithum, P., Kaewkungwal, J., Fungladda, W., Bussaratid, V., Koonsaeng, S., et al. (2009). Knowledge, attitudes, and acceptability of a human papillomavirus vaccine among healthcare providers. *Southeast Asian Journal of Tropical Medicine and Public Health*, 40(5), 1048.
- Sonntag, D. (2010). *Ontologies and Adaptivity in Dialogue for Question Answering*, volume 4. IOS Press.
- Sotiriadis, A., Dagklis, T., Siamanta, V., Chatzigeorgiou, K., Agorastos, T., Group, L. S., et al. (2012). Increasing fear of adverse effects drops intention to vaccinate after the introduction of prophylactic hpv vaccine. *Archives of gynecology and obstetrics*, 285(6), 1719–1724.
- Speer, R. & Havasi, C. (2013). ConceptNet 5: A Large Semantic Network for Relational Knowledge. In I. Gurevych & J. Kim (Eds.), *The People’s Web Meets NLP* (pp. 161–176). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Speer, R. & Lowry-Duda, J. (2017). Conceptnet at semeval-2017 task 2: Extending word embeddings with multilingual relational knowledge. *arXiv preprint arXiv:1704.03560*.

- Squid Ink (2019). Support icon. <http://www.iconarchive.com/show/free-flat-sample-icons-by-thesquid.ink/support-icon.html>. CC Attribution 3.0. Accessed 23 Apr 2019.
- Stamper, R., Liu, K., Hafkamp, M., & Ades, Y. (2000). Understanding the roles of signs and norms in organizations-a semiotic approach to information systems design. *Behaviour & Information Technology*, 19(1), 15–27.
- Steunebrink, B. R., Dastani, M., & Meyer, J.-J. C. (2009). The OCC model revisited. In *Proc. of the 4th Workshop on Emotion and Computing*.
- Stokley, S., Cohn, A., Dorell, C., Hariri, S., Yankey, D., Messonnier, N., & Wortley, P. M. (2011). Adolescent vaccination-coverage levels in the united states: 2006–2009. *Pediatrics*, (pp. peds–2011).
- Stretch, R., McCann, R., Roberts, S. A., Elton, P., Baxter, D., & Brabin, L. (2009). A qualitative study to assess school nurses' views on vaccinating 12–13 year old school girls against human papillomavirus without parental consent. *BMC Public Health*, 9(1), 1–6.
- Stupiansky, N. W., Alexander, A. B., & Zimet, G. D. (2012). Human papillomavirus vaccine and men: what are the obstacles and challenges? *Current opinion in infectious diseases*, 25(1), 86–91.
- Sundström, K., Tran, T. N., Lundholm, C., Young, C., Sparén, P., & Dahlström, L. A. (2010). Acceptability of hpv vaccination among young adults aged 18–30 years—a population based survey in sweden. *Vaccine*, 28(47), 7492–7500.

- Sussman, A. L., Helitzer, D., Sanders, M., Urquieta, B., Salvador, M., & Ndiaye, K. (2007). Hpv and cervical cancer prevention counseling with younger adolescents: implications for primary care. *The Annals of Family Medicine*, 5(4), 298–304.
- Swanson, J. (1992). What does it take to adapt instruction to the individual? a case study of one-to-one tutoring. *American Education Research Association, San Francisco, CA*.
- Tannen, D. (1982). Oral and literate strategies in spoken and written narratives. *Language*, (pp. 1–21).
- Tao, C., Wei, W.-Q., Solbrig, H. R., Savova, G., & Chute, C. G. (2010). Cntro: a semantic web ontology for temporal relation inferencing in clinical narratives. In *AMIA annual symposium proceedings*, volume 2010 (pp. 787).: American Medical Informatics Association.
- Tariq, S., Bhakta, S., Grimes, A., & Stevens, A. (2009). Attitudes of arkansas' primary care physicians toward the hpv vaccine. *The Journal of the Arkansas Medical Society*, 106(5), 113–117.
- Tartir, S. & Arpinar, I. B. (2007). Ontology evaluation and ranking using ontoqa. In *Semantic Computing, 2007. ICSC 2007. International Conference on* (pp. 185–192).: IEEE.
- Tartir, S., Arpinar, I. B., Moore, M., Sheth, A. P., & Aleman-Meza, B. (2005). Ontoqa: Metric-based ontology quality analysis. In *IEEE ICDM Workshop on Knowledge Acquisition from Distributed, Autonomous, Semantically Heterogeneous Data and Knowledge Sources*: IEEE.



- Thomas, T. L., Strickland, O., Diclemente, R., & Higgins, M. (2013). An opportunity for cancer prevention during preadolescence and adolescence: Stopping human papillomavirus (hpv)-related cancer through hpv vaccination. *Journal of Adolescent Health*, 52(5), S60–S68.
- Thompson, V. L. S., Arnold, L. D., & Notaro, S. R. (2012). African american parents' hpv vaccination intent and concerns. *Journal of health care for the poor and underserved*, 23(1), 290.
- Tielman, M. L., Neerincx, M. A., Bidarra, R., Kybartas, B., & Brinkman, W.-P. (2017). A Therapy System for Post-Traumatic Stress Disorder Using a Virtual Agent and Virtual Storytelling to Reconstruct Traumatic Memories. *Journal of Medical Systems*, 41(8).
- Tiro, J. A., Tsui, J., Bauer, H. M., Yamada, E., Kobrin, S., & Breen, N. (2012). Human papillomavirus vaccine use among adolescent girls and young adult women: an analysis of the 2007 california health interview survey. *Journal of Women's Health*, 21(6), 656–665.
- Toffolon-Weiss, M., Hagan, K., Leston, J., Peterson, L., Provost, E., & Hennessy, T. (2008). Alaska native parental attitudes on cervical cancer, hpv and the hpv vaccine. *International journal of circumpolar health*, 67(4).
- Tolk, A., Diallo, S. Y., & Padilla, J. J. (2012). Semiotics, entropy, and interoperability of simulation systems: mathematical foundations of m&s standardization. In *Proceedings of the Winter Simulation Conference* (pp. 243): Winter Simulation Conference.

- Tozzi, A. E., Ravà, L., Stat, D., Pandolfi, E., Marino, M. G., & Ugazio, A. G. (2009). Attitudes towards hpv immunization of italian mothers of adolescent girls and potential role of health professionals in the immunization program. *Vaccine*, 27(19), 2625–2629.
- U.S. Cancer Statistics Working Group (2013). *United States Cancer Statistics: 1999–2009 Incidence and Mortality Web-based Report*. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute.
- US Department of Health and Human Services and Office of Disease Prevention and Health Promotion and others (2012). Healthy people 2020.
- US Food and Drug Administration (2018). Fda approves expanded use of gardasil 9 to include individuals 27 through 45 years old. Retrieved from: <https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm622715.htm>.
- US Food and Drug Administration and others (2006). Fda licenses new vaccine for prevention of cervical cancer and other diseases in females caused by human papillomavirus. *FDA News*, 10.
- Vadaparampil, S. T., Kahn, J. A., Salmon, D., Lee, J.-H., Quinn, G. P., Roetzheim, R., Bruder, K., Malo, T. L., Proveaux, T., Zhao, X., et al. (2011). Missed clinical opportunities: provider recommendations for hpv vaccination for 11–12 year old girls are limited. *Vaccine*, 29(47), 8634–8641.

- Valentino, K. & Poronsky, C. B. (2015). Human papillomavirus infection and vaccination. *Journal of pediatric nursing*.
- Vamos, C. A., McDermott, R. J., & Daley, E. M. (2008). The hpv vaccine: framing the arguments for and against mandatory vaccination of all middle school girls. *Journal of School Health*, 78(6), 302–309.
- Van Mulken, S., André, E., & Müller, J. (1999). An empirical study on the trustworthiness of life-like interface agents. In *HCI (2)*(pp. 152–156).: Citeseer.
- VanLehn, K., Graesser, A. C., Jackson, G. T., Jordan, P. W., Olney, A., & Rosé, C. P. (2007). When are tutorial dialogues more effective than reading? *Cognitive Science*, 31(1), 3–62.
- Velicer, W. F., Prochaska, J. O., Fava, J. L., Laforge, R. G., & Rossi, J. S. (1999). Interactive versus noninteractive interventions and dose-response relationships for stage-matched smoking cessation programs in a managed care setting. *Health psychology : official journal of the Division of Health Psychology, American Psychological Association*, 18(1), 21.
- Verhoeven, V., Baay, M., Colliers, A., Verster, A., Van Royen, P., Avonts, D., & Vermorken, J. B. (2006). The male factor in cervical carcinogenesis: a questionnaire study of men's awareness in primary care. *Preventive medicine*, 43(5), 389–393.
- Villa, L. L., Perez, G., Kjaer, S. K., Paavonen, J., Lehtinen, M., Munoz, N., Sigurdsson, K., Hernandez-Avila, M., Skjeldestad, F. E., Thoresen, S., et al. (2007). Quadrivalent vaccine

- against human papillomavirus to prevent high-grade cervical lesions. *New England Journal of Medicine*, 356(19), 1915–1927.
- Voorhees, E. M. & Tice, D. M. (2000). Building a question answering test collection. In *Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 200–207).: ACM.
- Vrandečić, D. (2009). Handbook on ontologies. by *A. Staab and R. Studer. Springer Berlin Heidelberg. Chap. Ontology Evaluation*, (pp. 293–313).
- Walker, T. Y., Elam-Evans, L. D., Singleton, J. A., Yankey, D., Markowitz, L. E., Fredua, B., Williams, C. L., Meyer, S. A., & Stokley, S. (2017). National, regional, state, and selected local area vaccination coverage among adolescents aged 13–17 years—united states, 2016. *MMWR. Morbidity and mortality weekly report*, 66(33), 874.
- Waller, J., Marlow, L. A., & Wardle, J. (2006). Mothers’ attitudes towards preventing cervical cancer through human papillomavirus vaccination: a qualitative study. *Cancer Epidemiology Biomarkers & Prevention*, 15(7), 1257–1261.
- Wang, C., Xiong, M., Zhou, Q., & Yu, Y. (2007). Panto: A portable natural language interface to ontologies. In *The Semantic Web: Research and Applications* (pp. 473–487). Springer.
- Wang, D., Cunningham, R. M., Boom, J., Amith, M., & Tao, C. (2016). Towards a HPV Vaccine Knowledgebase For Patient Education Content. *Studies in Health Technology and Informatics*.

- Watkins, K. L., Reitzel, L. R., Wetter, D. W., & McNeill, L. H. (2015). Hpv awareness, knowledge and attitudes among older african-american women. *American journal of health behavior*, 39(2), 205–211.
- Watson, D. & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological bulletin*, 98(2), 219.
- Watson, M., Saraiya, M., Ahmed, F., Cardinez, C. J., Reichman, M. E., Weir, H. K., & Richards, T. B. (2008a). Using population-based cancer registry data to assess the burden of human papillomavirus-associated cancers in the united states: Overview of methods. *Cancer*, 113(S10), 2841–2854.
- Watson, M., Saraiya, M., Benard, V., Coughlin, S. S., Flowers, L., Cokkinides, V., Schwenn, M., Huang, Y., & Giuliano, A. (2008b). Burden of cervical cancer in the united states, 1998–2003. *Cancer*, 113(S10), 2855–2864.
- Wei, F., Moore, P. C., & Green, A. L. (2013). Geographic variability in human papillomavirus vaccination among us young women. *American journal of preventive medicine*, 44(2), 154–157.
- Weinstock, H., Berman, S., & Cates Jr, W. (2004). Sexually transmitted diseases among american youth: incidence and prevalence estimates, 2000. *Perspectives on sexual and reproductive health*, (pp. 6–10).

- Weisberg, E., Bateson, D., Skinner, S., McCaffery, K., et al. (2009). Hpv vaccination catch up program: utilisation by young australian women. *Australian family physician*, 38(1-2), 72.
- Weiss, T. W., Rosenthal, S. L., & Zimet, G. D. (2011). Attitudes toward hpv vaccination among women aged 27 to 45. *ISRN obstetrics and gynecology*, 2011.
- White, J. (2015). Robert kennedy jr. warns of vaccine-linked 'holocaust'.  
<http://www.sacbee.com/news/politics-government/capitol-alert/article17814440.html>.
- Williams, K., Forster, A., Marlow, L., & Waller, J. (2011). Attitudes towards human papillomavirus vaccination: a qualitative study of vaccinated and unvaccinated girls aged 17–18 years. *Journal of Family Planning and Reproductive Health Care*, 37(1), 22–25.
- Williams, W. W., Lu, P.-J., O'Halloran, A., Bridges, C. B., Pilishvili, T., Hales, C. M., & Markowitz, L. E. (2014). Noninfluenza vaccination coverage among adults-united states, 2012. *MMWR. Morbidity and mortality weekly report*, 63(5), 95–102.
- Williams, W. W., Lu, P.-J., Saraiya, M., Yankey, D., Dorell, C., Rodriguez, J. L., Kepka, D., & Markowitz, L. E. (2013). Factors associated with human papillomavirus vaccination among young adult women in the united states. *Vaccine*, 31(28), 2937–2946.
- Wong, L. P. (2008). Young multiethnic women's attitudes toward the hpv vaccine and hpv vaccination. *International Journal of Gynecology & Obstetrics*, 103(2), 131–135.
- Wong, L. P. (2009). Physicians' experiences with hpv vaccine delivery: evidence from developing country with multiethnic populations. *Vaccine*, 27(10), 1622–1627.

- Wong, L. P. & Sam, I.-C. (2010). Ethnically diverse female university students' knowledge and attitudes toward human papillomavirus (hpv), hpv vaccination and cervical cancer. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 148(1), 90–95.
- Wood, D., Wood, H., & Middleton, D. (1978). An experimental evaluation of four face-to-face teaching strategies. *International journal of behavioral development*, 1(2), 131–147.
- Woodhall, S. C., Lehtinen, M., Verho, T., Huhtala, H., Hokkanen, M., & Kosunen, E. (2007). Anticipated acceptance of hpv vaccination at the baseline of implementation: a survey of parental and adolescent knowledge and attitudes in finland. *Journal of Adolescent Health*, 40(5), 466–469.
- Wooldridge, M. (2009). *An introduction to multiagent systems*. John Wiley & Sons.
- Wooldridge, M. & Jennings, N. R. (1995). Intelligent agents: Theory and practice. *The knowledge engineering review*, 10(2), 115–152.
- Yeganeh, N., Curtis, D., & Kuo, A. (2010). Factors influencing hpv vaccination status in a latino population; and parental attitudes towards vaccine mandates. *Vaccine*, 28(25), 4186–4191.
- Ylitalo, K. R., Lee, H., & Mehta, N. K. (2013). Health care provider recommendation, human papillomavirus vaccination, and race/ethnicity in the us national immunization survey. *American journal of public health*, 103(1), 164–169.

- Young, D., Xiao, C. C., Murphy, B., Moore, M., Fakhry, C., & Day, T. A. (2015). Increase in head and neck cancer in younger patients due to human papillomavirus (hvp). *Oral oncology*, 51(8), 727–730.
- Young, M., Sparrow, D., Gottlieb, D., Selim, A., & Friedman, R. (2001). A telephone-linked computer system for copd care. *CHEST Journal*, 119(5), 1565–1575.
- Yu, A. C. (2006). Methods in biomedical ontology. *Journal of Biomedical Informatics*, 39(3), 252–266.
- Yu, L. (2011). *A Developer's Guide to the Semantic Web*. Heidelberg; New York: Springer, 2011 edition edition.
- Zak-Place, J. & Stern, M. (2004). Health belief factors and dispositional optimism as predictors of std and hiv preventive behavior. *Journal of American College Health*, 52(5), 229–236.
- Zhou, L. & Zhang, D. (2004). Building a misinformation ontology. In *Web Intelligence, 2004. WI 2004. Proceedings. IEEE/WIC/ACM International Conference on* (pp. 445–448): IEEE.
- Zhou, L. & Zhang, D. (2007). An ontology-supported misinformation model: Toward a digital misinformation library. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 37(5), 804–813.
- Zhou, Y., Freedman, R., Glass, M., Michael, J. A., Rovick, A. A., & Evens, M. W. (1999). Delivering hints in a dialogue-based intelligent tutoring system. In *AAAI/IAAI* (pp. 128–134).



- Zhu, H., Liu, D., Bayley, I., Aldea, A., Yang, Y., & Chen, Y. (2017). Quality model and metrics of ontology for semantic descriptions of web services. *Tsinghua Science and Technology*, 22(3), 254–272.
- Zimet, G. D. (2005). Improving adolescent health: focus on hpv vaccine acceptance. *Journal of Adolescent Health*, 37(6), S17–S23.
- Zimet, G. D., Rosberger, Z., Fisher, W. A., Perez, S., & Stupiansky, N. W. (2013). Beliefs, behaviors and hpv vaccine: correcting the myths and the misinformation. *Preventive medicine*, 57(5), 414–418.
- Zimet, G. D., Weiss, T. W., Rosenthal, S. L., Good, M. B., & Vichnin, M. D. (2010). Reasons for non-vaccination against hpv and future vaccination intentions among 19-26 year-old women. *BMC women's health*, 10(1), 27.