Natural Language Dialogue Systems using AI and ML

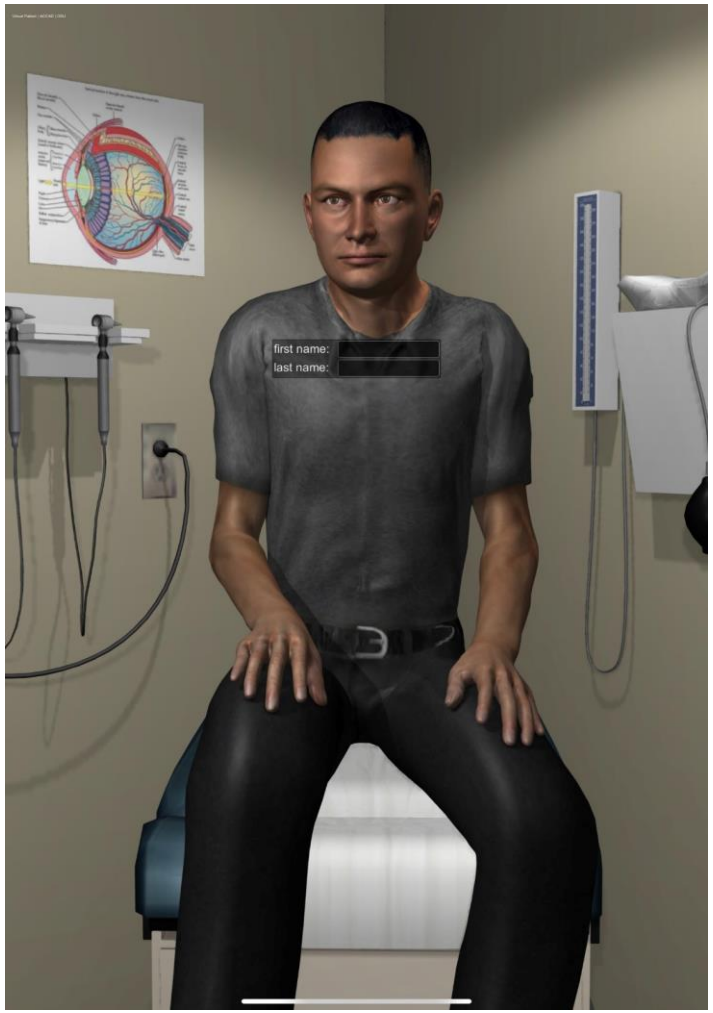
Douglas Danforth, Ph.D.
Associate Professor
Department of Obstetrics and Gynecology

Wexner Medical Center

Improving People's Lives
through innovation in research, education and patient care

virtual

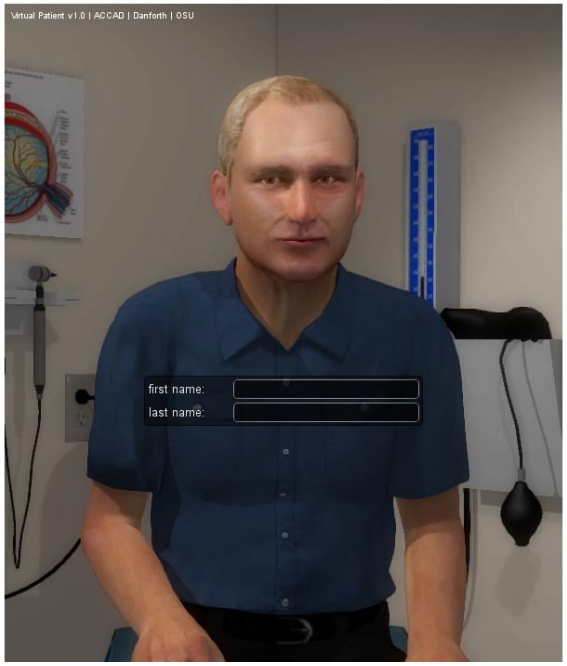
Virtual Patient



- Artificial Intelligence - simulate “real” patient
- Conversational – can understand and respond to student questions
- Provide immediate feedback – students can practice multiple times

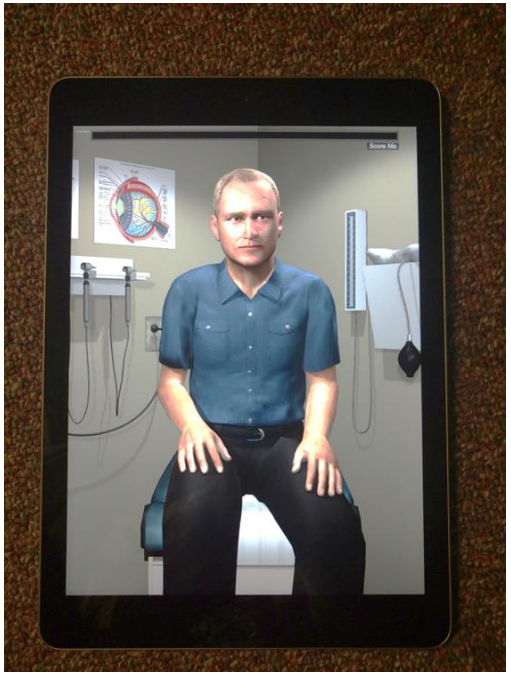
Virtual Patient

Web-based
chatted conversation



Clinical Skills Center
Spoken conversation
(Dragon)

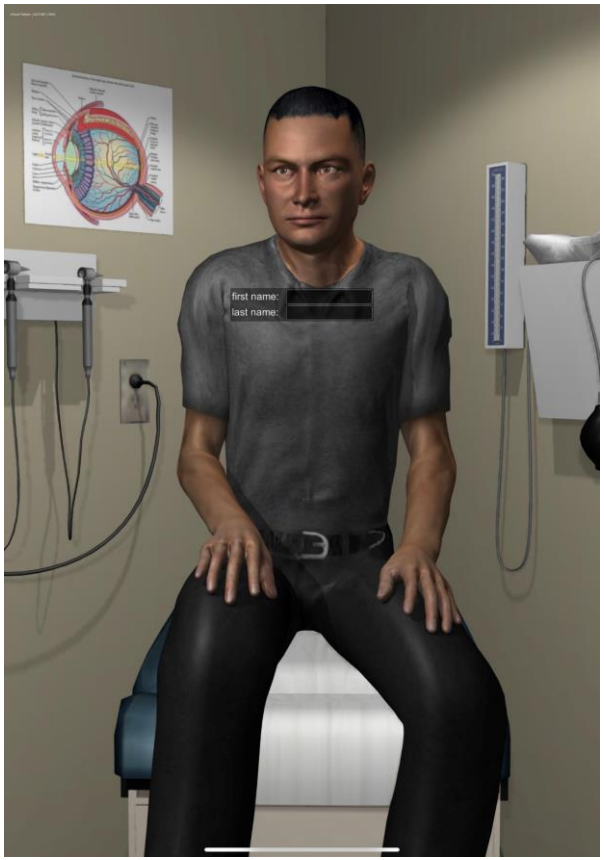
iPad
Spoken conversation
(Watson)



Wexner Medical Center

Natural Language Understanding

- Input type
  - Typed or spoken
- Question identification
  - Rule based vs ML/NLU
- Output
  - TTS vs pre-recorded



Wexner Medical Center

Dialogue Management

Input type

Chatted (typed)

- Questions are more succinct
- Conversations shorter
- Typos considerable

- *“Waht brings you in today?”*
- *“When did hte pain start?”*
- *“Anything besieds the back pain?”*
- *“Tell me more.”*

Spoken (STT)

- Questions are much longer
- Multiple questions in one query
- Conversations longer
- Typos not important
- STT mistranslations important

- *“So I'd be probing seems to help so that's good and so you started you said this began about 4 to 5 years ago you had another episode.”*
- *“Are you generally happy with your life now? your job? house?”*

Wexner Medical Center

Question Identification

Rule based (pattern matching)

- Based on NLP Software ChatScript
- Initial NLP Processing (spell checking, canonization, POS tagging, etc)
- State aware – pronoun resolution
- ~3,000 rules to manage dialogue

ChatScript

Dialogue organized by History of Present Illness, Past Medical History, Family History, Social History

```
#! What brings you in today
u: What_brings_you_in_today (["what brings you" "why are you"] in today)  ^factanswer("I'm hoping you can help me out with a problem I've been having.")
u: (![when other else] [why what] * ["do for" bring help come] * [today tonight evening afternoon morning day]) ^reuse (What_brings_you_in_today)
u: (![when other else] "what brings you in") ^reuse (What_brings_you_in_today)
u: (![when other else] "what brings you in today") ^reuse (What_brings_you_in_today)
u: (![when other else] "what brings you here") ^reuse (What_brings_you_in_today)

#! Any other problems
u: Any_other_problems1 (any other problems)  ^factanswer("I don't believe so.")
u: (!health !medical << [other else] [symptom issue concern problem "bothering you" discuss] >>) ^reuse (Any_other_problems1)
u: (!try anything else) ^reuse (Any_other_problems1)

#! Are you taking any medication for the 1 $chiefcomplaint=1
u: Taking_medication_for_the_1 (are you taking any medication for the $chiefcomplaint?)  ^factanswer("I don't take any medications.")
u: (![how long other] [use take on try] * [alleviate get_rid_of reduce diminish "help with" for] *~2 [~itwords $chiefcomplaint?]) ^reuse (Taking_medication_for_the_1)
u: (![how long other] [use take on try "no"] * ~medicines * [~itwords $chiefcomplaint?]) ^reuse (Taking_medication_for_the_1)
```

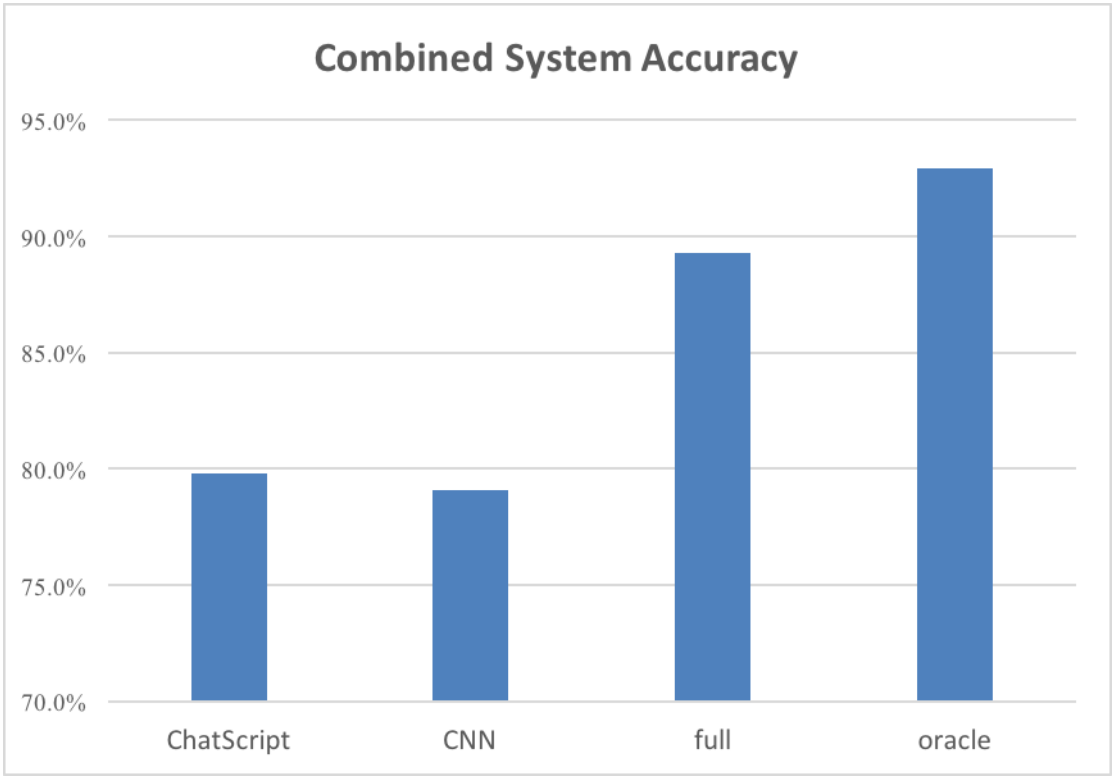
Wexner Medical Center

Dialogue Accuracy

Dialogue Type	Typed
Type of Case	Simple
Type of History	Complete
Students	First Year (n=21)
Total questions asked	1396
Answered correctly	83%
Answered incorrectly	6.4%
Not answered	10%

Dialogue accuracy improvements with ML/NLU

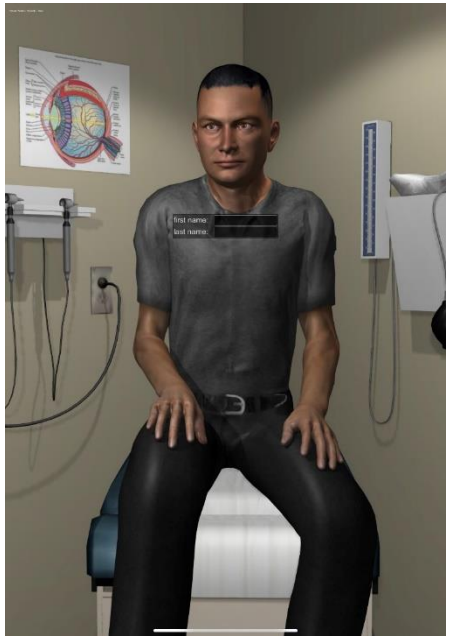
- Train Convolutional Neural Network analyzer (CNN)
- Used word - and character-based CNNs to identify and classify input questions
- Binary classifier to choose between CNN and ChatScript



Wexner Medical Center

Fall 2018 – Mr. Carlos Martinez

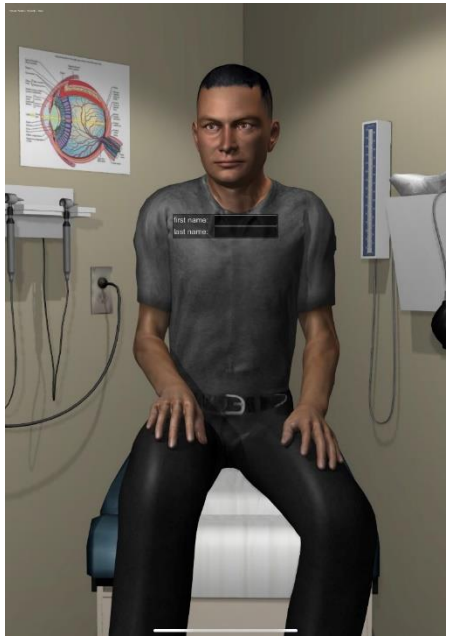
- Practice history taking by first year medical students
- HOPI, PMH, Family History, Social History
- ~200 Conversations
- > 12,000 Questions



	All inputs	Watson Ex (71.4%)	Watson Acceptable (89.5%)	Watson Gibberish (10.5%)
ChatScript Correct	73.5%	78.3%	76.0%	50.2%
CNN Correct	72.7%	74%	73.6%	59.6%
Acceptable Response	76.2%	80.8%	78.7%	53.8%
Question Not Understd	16.5%	11%	11.7%	35.4%

Fall 2018 – Mr. Carlos Martinez

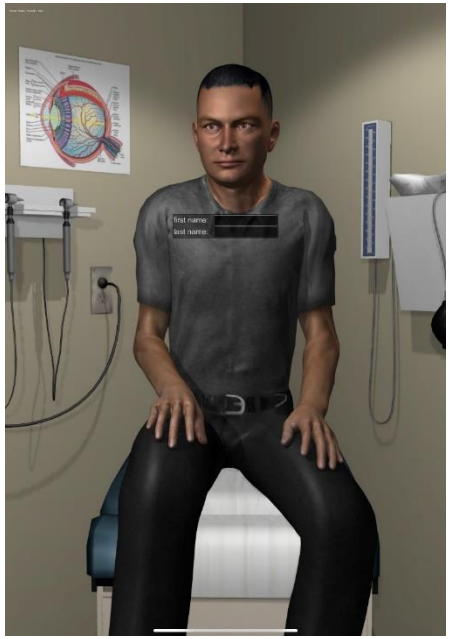
- Practice history taking by first year medical students
- HOPI, PMH, Family History, Social History
- ~200 Conversations
- > 12,000 Questions



	All inputs	Watson Exact (71.4%)	Watson Acceptable (89.5%)	Watson Gibberish (10.5%)

Fall 2018 – Mr. Carlos Martinez

- Practice history taking by first year medical students
- HOPI, PMH, Family History, Social History
- ~200 Conversations
- > 12,000 Questions



	All inputs	Watson Exact (71.4%)	Watson Acceptable (89.5%)	Watson Gibberish (10.5%)

Word Error Rate					
Microsoft	Google	Watson untrained	Watson Acoustic	Watson Language	Watson Both
0.078	0.053	0.079	0.069	0.058	0.056

Natural Language Understanding

Text-to-Speech

- “Voices” limited
- Quality varies
- Limited inflection
- Somewhat slow

Pre-recorded

- Time-consuming to create
- More difficult to deliver
- Difficult to update or modify
- More realistic

Summary

- Using Natural Language Understanding, Virtual Patients can understand, respond, categorize, and “assess” student performance, enabling students to practice their history taking skills and receive immediate feedback.

- Dialogue management can use fairly simply rule based approaches or more sophisticated NLU systems.
  - Rule based systems require hand crafting rules but can work with limited data.
  - NLU systems can be more robust but require significant training data.
  - Hybrid systems can maximize accuracy.

- Depending on the application, modern cloud-based systems may be easier to implement.

Support

Artificial Intelligence based Virtual Reality (VR) Simulation of Provider-patient interaction to enhance cultural competency using a simulated patient with limited English proficiency. Medicaid Equity Simulation Project $455,103 2018 – 2019

Virtual Patients for Medication Assisted Treatment of Opioid Use Disorder. DHHS Health Resources and Services Administration (Supplement) $100,000 2018 – 2019

Using Automatically Generated Paraphrases and Discriminative ASR Training to Author Robust Question-Answering Dialogue Systems. National Science Foundation $499,904 2016 – 2019

Creating the Complete Virtual Standardized Patient: Integrating Natural Language Ability into Clinical Reasoning Cases to Assess Information Gathering and Clinical Reasoning. The Institute for Innovative Technologies in Medical Education/Med-U $24,250, 2015 – 2019

Virtual Patient Simulations to Assess Data Gathering and Clinical Reasoning National Board of Medical Examiners Edward J. Stemmler Education Research Fund, $149,862, 2012 – 2014

Virtual Patients in the 2012 Curriculum The Ohio State University College of Medicine Innovation Fund, $45,000 2011- 2013

Virtual Reality: A Unique Means to Teach the Reality of the Patient-Centered Medical Home. DHHS Health Resources and Services Administration, $ 1,465,655 2010 – 2015

Development of Virtual Patients using the Second Life platform. Perinatal Resources Inc, $30,000 2008-2011