Low resourced and long tailed conversational AI systems

Eric Fosler-Lussier
The Ohio State University
Symposium on NLP and Simulation, October 2022

with Doug Danforth, Mike White, William Schuler, Kellen Maicher, Laura Zimmerman, Ayaz Hyder, Huan Sun
Adam Stiff, Prashant Serai, Vishal Sunder, Evan Jaffe, Lifeng Jin, David King, Amad Hussain, Harsh Pandey, Erika Goetz

Work partially funded by NSF IIS-1618336, NSF OAC-2112606 (ICICLE NSF AI Institute)

This is a story of partnerships

Doug Danforth

THE OHIO STATE UNIVERSITY
COLLEGE OF MEDICINE

This is a story of partnerships

Doug Danforth

THE OHIO STATE UNIVERSITY
COLLEGE OF MEDICINE

200 medical students/class

This is a story of partnerships

Doug Danforth

THE OHIO STATE UNIVERSITY
COLLEGE OF MEDICINE

200 medical students/class

Students are evaluated in taking medical histories and differential diagnosis with **standardized patients**

This is a story of partnerships

Doug Danforth

THE OHIO STATE UNIVERSITY
COLLEGE OF MEDICINE

200 medical students/class

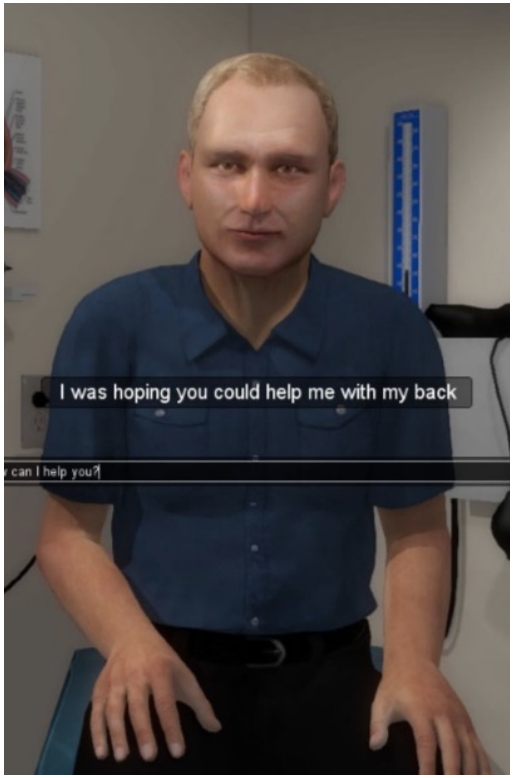
Brillig Understanding, Inc.

Why GitHub?

ChatScript / **ChatScript**

THE OHIO STATE UNIVERSITY
COLLEGE OF ARTS AND SCIENCES

ADVANCED COMPUTING CENTER
FOR THE **ARTS** AND **DESIGN**



Students **can practice** taking medical histories and differential diagnosis with **virtual standardized patients**

This is a story of partnerships

Doug Danforth

THE OHIO STATE UNIVERSITY
COLLEGE OF MEDICINE

200 medical students/class

Knowledge / Rule Patterns

Brillig Understanding, Inc.

Why GitHub?

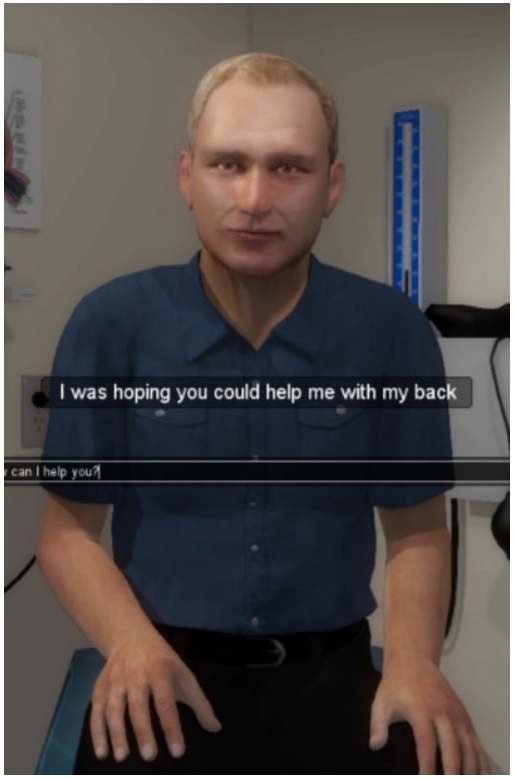
ChatScript / **ChatScript**

Dialog Engine

THE OHIO STATE UNIVERSITY
COLLEGE OF ARTS AND SCIENCES

ADVANCED COMPUTING CENTER
FOR THE **ARTS** AND **DESIGN**

Avatar



Students **can practice** taking medical histories and differential diagnosis with **virtual standardized patients**

This is a story of partnerships

Doug Danforth

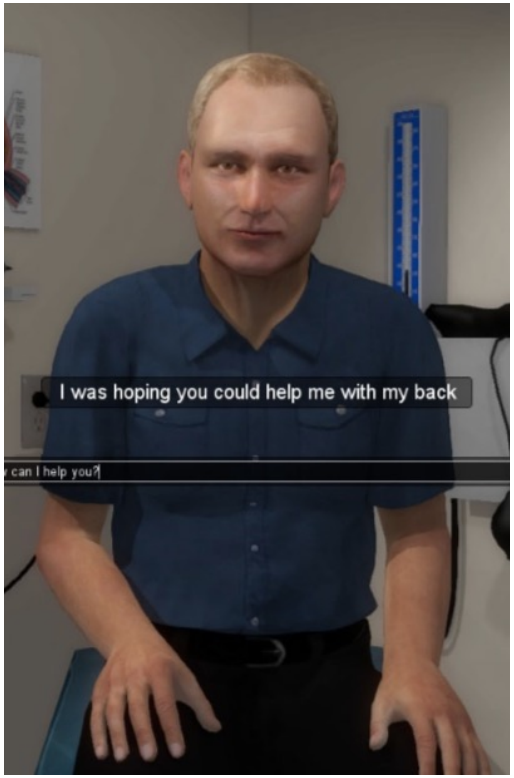
THE OHIO STATE UNIVERSITY
COLLEGE OF MEDICINE

Doug approaches the NLP folks in OSU CSE/Linguistics with questions:

How do we make understanding systems more robust?

How do we convert a chatbot into speech enabled system?

We said: sure, we can!

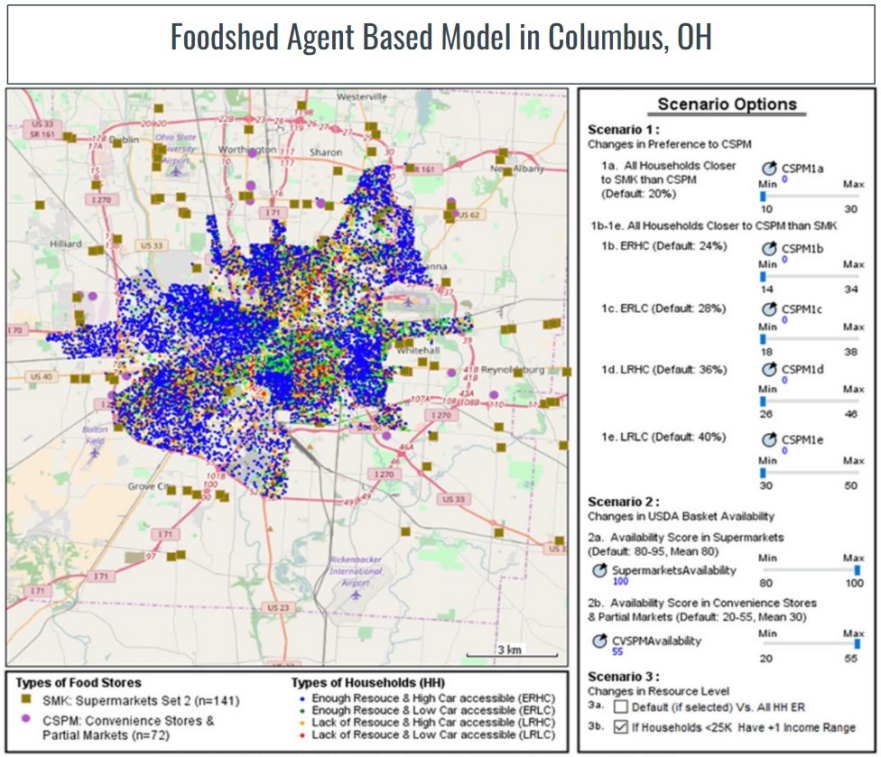


Students **can practice** taking medical histories and differential diagnosis with **virtual standardized patients**

This is a story of partnerships

Ayaz Hyder

ICICLE
DEMOCRATIZING AI



Agent based model can answer “what if” questions like “what happens to the people around the Kroger at 5th and Neil Ave if it closes?”

Question: can we build front ends that allow politicians and decisionmakers to query the model naturally?

Some lessons learned through partnership

NLP folks' perspective can be very different than domain researchers building systems for a task

- We both want to make systems work well
  - In the lab, we care about measuring performance
  - Our partners want to fix things when there is a bug
  - This difference can change course of research

Leads to a hybrid approach:
Need to blend rule-based approaches with neural approaches

- For these tasks, it's important that the answers are exact
  - Doctors need quality control over answers in VP
  - Grocery store model needs to control output
  - This is a very different than our typical systems that use NL generation
  - Requires a significant level of expert development

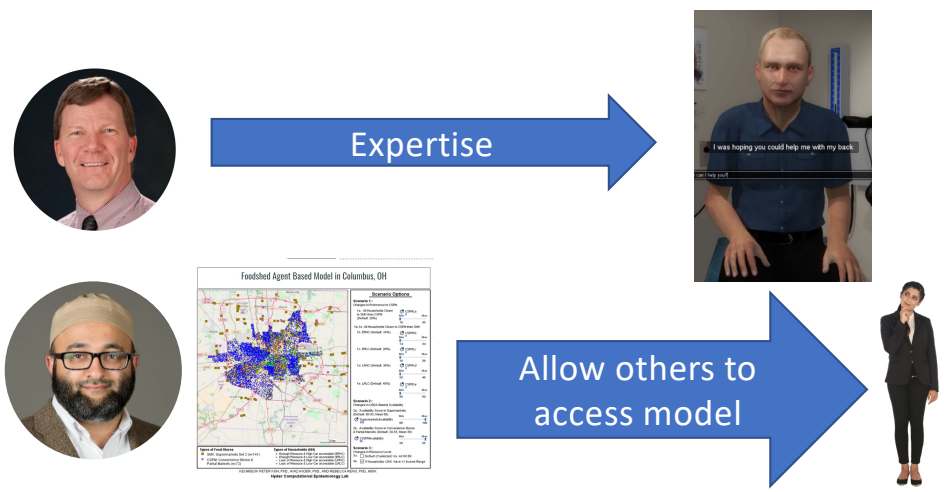
Design of Conversational AI system needs to be easily Influenced by domain experts

Low Resourced and Long Tailed Conversational AI

Low resource:

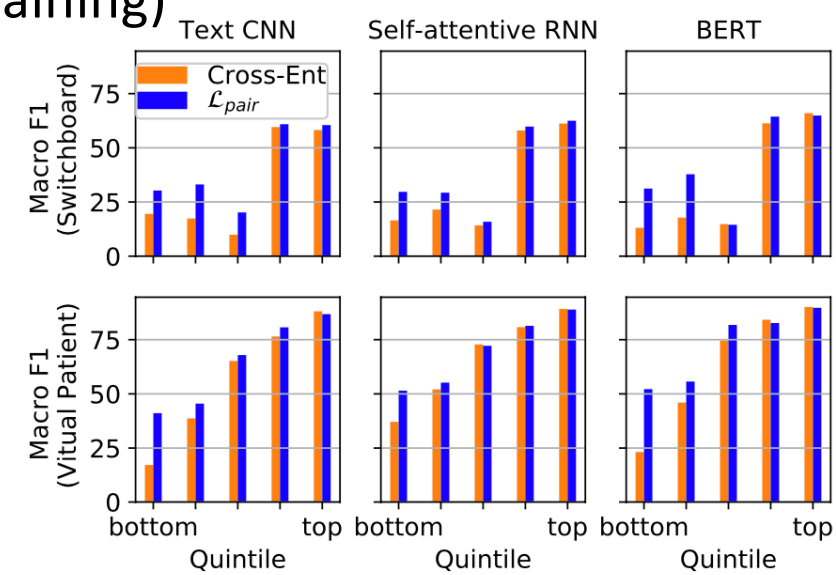
Can be hard to collect data; need to leverage other tasks.

Expert time is expensive!



Long tailed:

Many of the functions / classes might be rare (see them once during training)



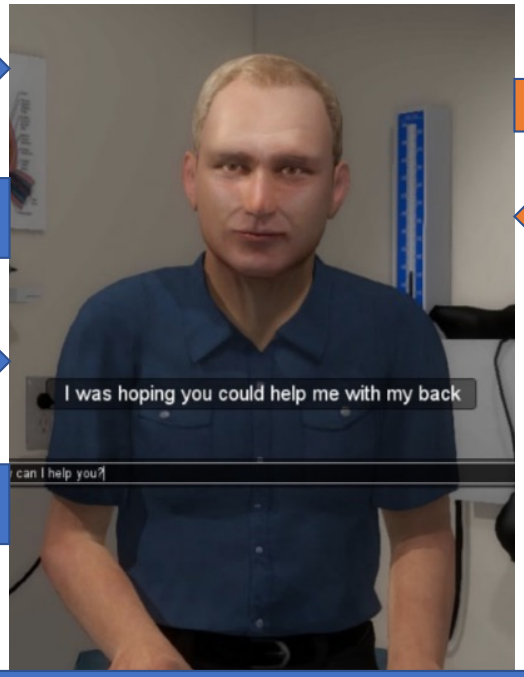
Virtual Patient: The Core Problem

How can I help you?

I was hoping you could help me with my back

Are you feeling pain?

Yes, it has been hurting for a while



Query

Canned Response

350+ Question Types

One or more patterns per type

Growing database of question examples

Significant # of singletons, sometimes new types

Report

4 of 23 categories covered…

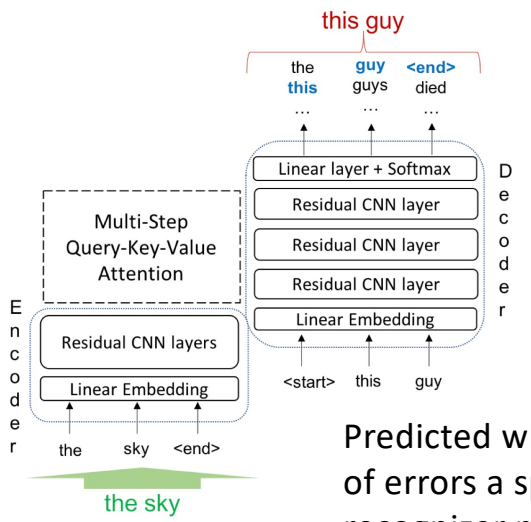
Data for training neural QA systems is sparse – long tail, low resource
Many classes, but some classes have few (or zero) examples
Initially: no speech data, only text data

Hello, oreo.

Hello, how are you?

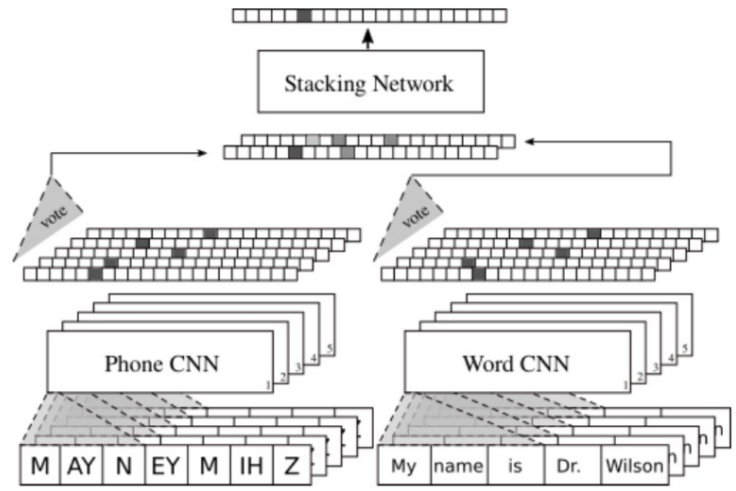
Problem 1: Speechifying a chatbot

Problem: Virtual patient had a typed interface, domain experts wanted a spoken interface. No in-domain speech data for training!



Serai, Wang, FL ICASSP 2019;
Serai, Stiff, FL ICASSP 2020

Predicted what kinds of errors a speech recognizer might make on this data, used it to augment training data.



Stiff, Serai, FL ICASSP 2019

Treat mispronunciations like “spoken typos”, use phonetic representations to increase robustness.

Problem 2: Handling the long tail

- Virtual Patient can answer 350+ classes of answers
  - Some are very frequent: How are you today?
  - Some are infrequent: Are you nervous?
- Combined rule-based system and neural system works better

- How do we improve trained systems?
  - Attention-based models that focus on extracting parts of rare classes
  - Few-shot learning through pairwise training

Self attentive RNN model

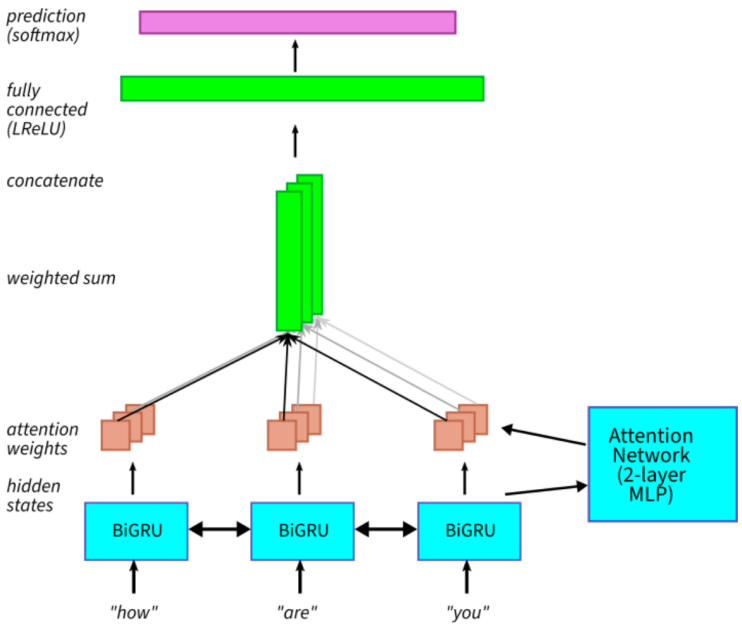


Figure 5.1: The self-attentive RNN model.

Rare classes are sometimes compositions of features from more common classes.

Use *attention* to help combine representations: this seems to help with rare classes.

Stiff & FL SigDial 2020

Attention improves rare class prediction

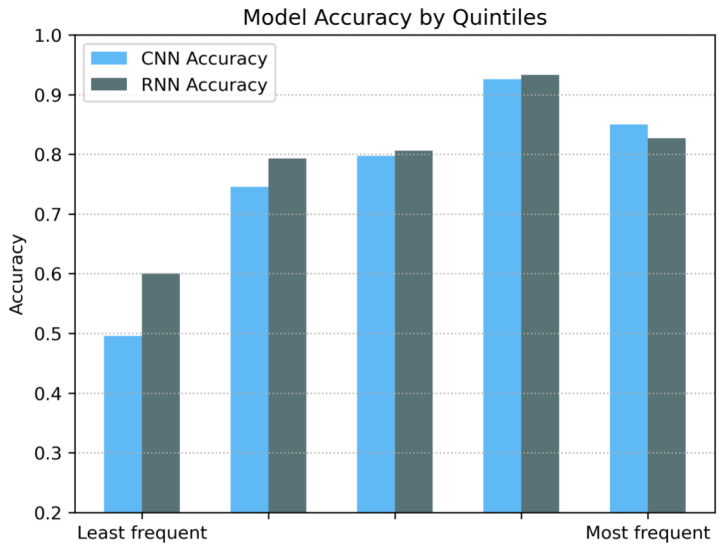


Figure 5.2: Quintile accuracies for the tested RNN and CNN baseline

Stiff & FL SigDial 2020

Changing how we train can also help

Rather than training a direct classifier, create nearest neighbor classifier that is informed by classification: pairwise mixup training

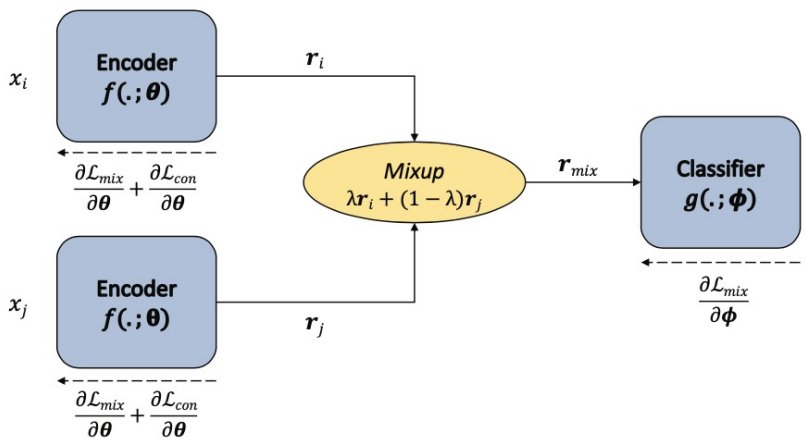


Fig. 1: Model overview. Sentences x_i and x_j are fed into the same encoder and interpolated thereafter. Solid lines indicate forward-propagation. Dashed lines indicate gradient flow.

Sunder & FL ICASSP 2021

Pairwise mixup training does two things:

1) Makes representations of sentences of the same class closer together; different classes further apart.
2) Requires the system to create representations that can be disentangled by mixing them during training (Fig 1)

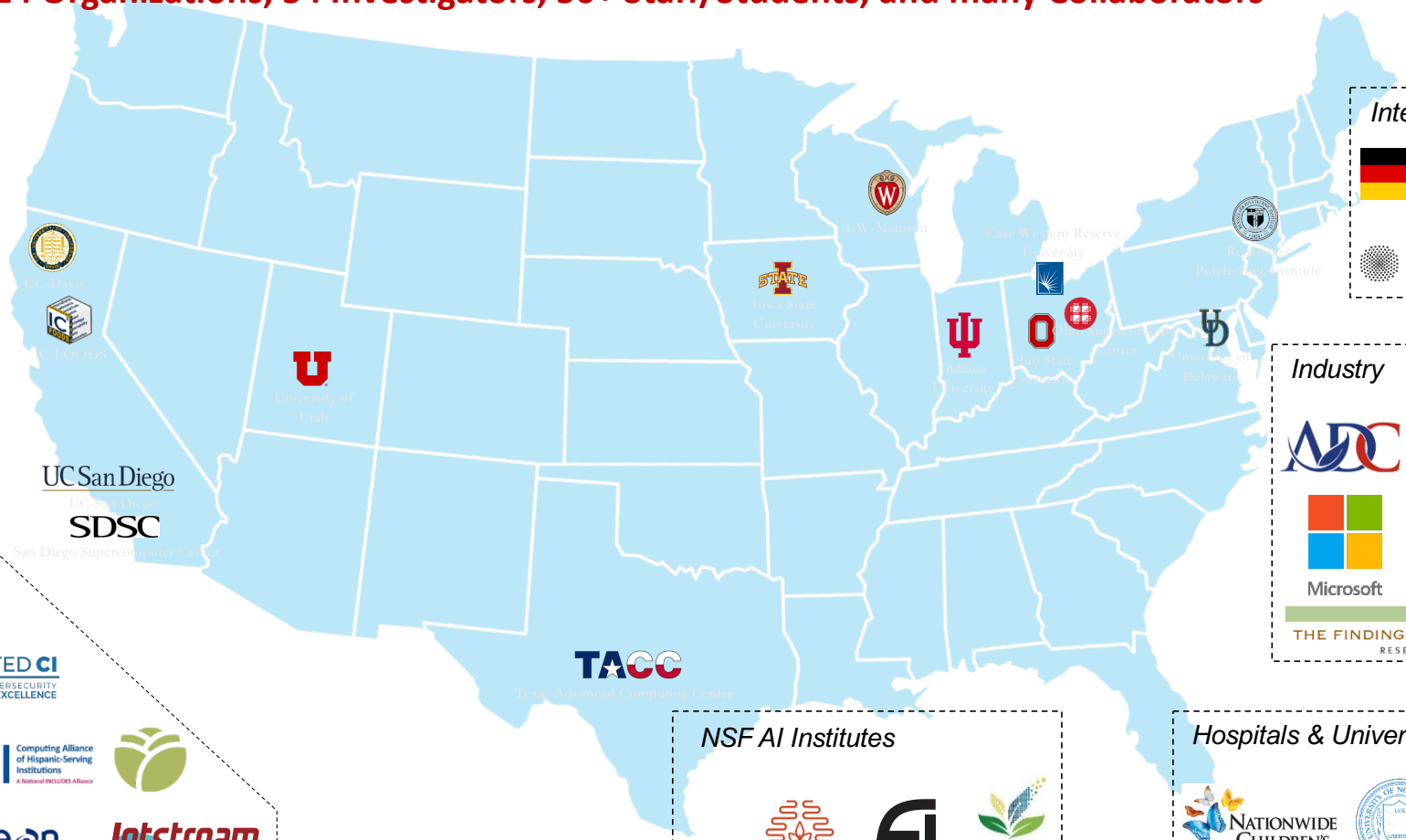
ICICLE: Intelligent CyberInfrastructure with Computational Learning in the Environment

- "Democratizing AI": how do we bring AI capabilities from HPC center to edge devices to domain scientists working in the field?
  - Digital Agriculture (e.g. drones for crop monitoring)
  - Wildlife Ecology (e.g. camera traps for monitoring wildlife)
  - Smart Foodsheds (e.g. modeling effects of grocery store placements)

http://icicle.ai

ICICLE NSF AI Institute Participation:

14 Organizations, 54 Investigators, 50+ Staff/Students, and many Collaborators



Govt. Agencies & National Labs

- Lawrence Livermore National Laboratory
- U.S. Fish & Wildlife Service
- Brookhaven National Laboratory
- Fermilab

Research Institutes

- National Center for Women & Information Technology
- MolSSI
- SGCI Science Gateways Community Institute
- Trusted CI The NSF Cybersecurity Center of Excellence
- ECEP
- CAHSI Computing Alliance of Hispanic-Serving Institutions
- SCEC An NSF+USGS Center
- NSF
- neon Operated by Battelle
- Jetstream

International

- Max-Planck-Gesellschaft
- University of Stuttgart Germany

Industry

- CNH Industrial
- ADC
- Wild Me
- Microsoft
- Farm 2 Facts
- The Findings Group, LLC Research & Evaluation

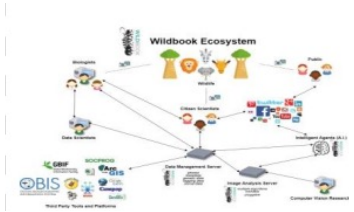
NSF AI Institutes

- AI Institute Food Systems (AIFS)

Hospitals & Universities

- Nationwide Children's
- University of North Carolina Chapel Hill
- Princeton

Objectives: Intelligent Cyberinfrastructure for the Computing Continuum



Animal Ecology

Digital Agriculture

Smart Foodsheds

ICICLE: Intelligent CyberInfrastructure with Computational Learning in the Environment

Systems AI Foundational Research for CI

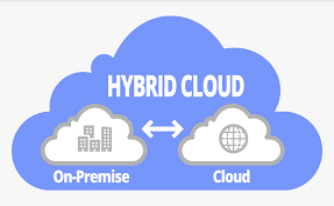
Intelligent Cyber Infrastructure

CI for AI

AI for "CI for AI"

On Field Sensors

Edge & Near Edge



Clouds

HPC Systems & Data Centers

Integrating a broad range of

- Scientists-in-the-field
- Engineers
- Educators
- Collaborative partners
- Institutions

under one roof enables

democratized,

adaptable,

plug-and-play AI,

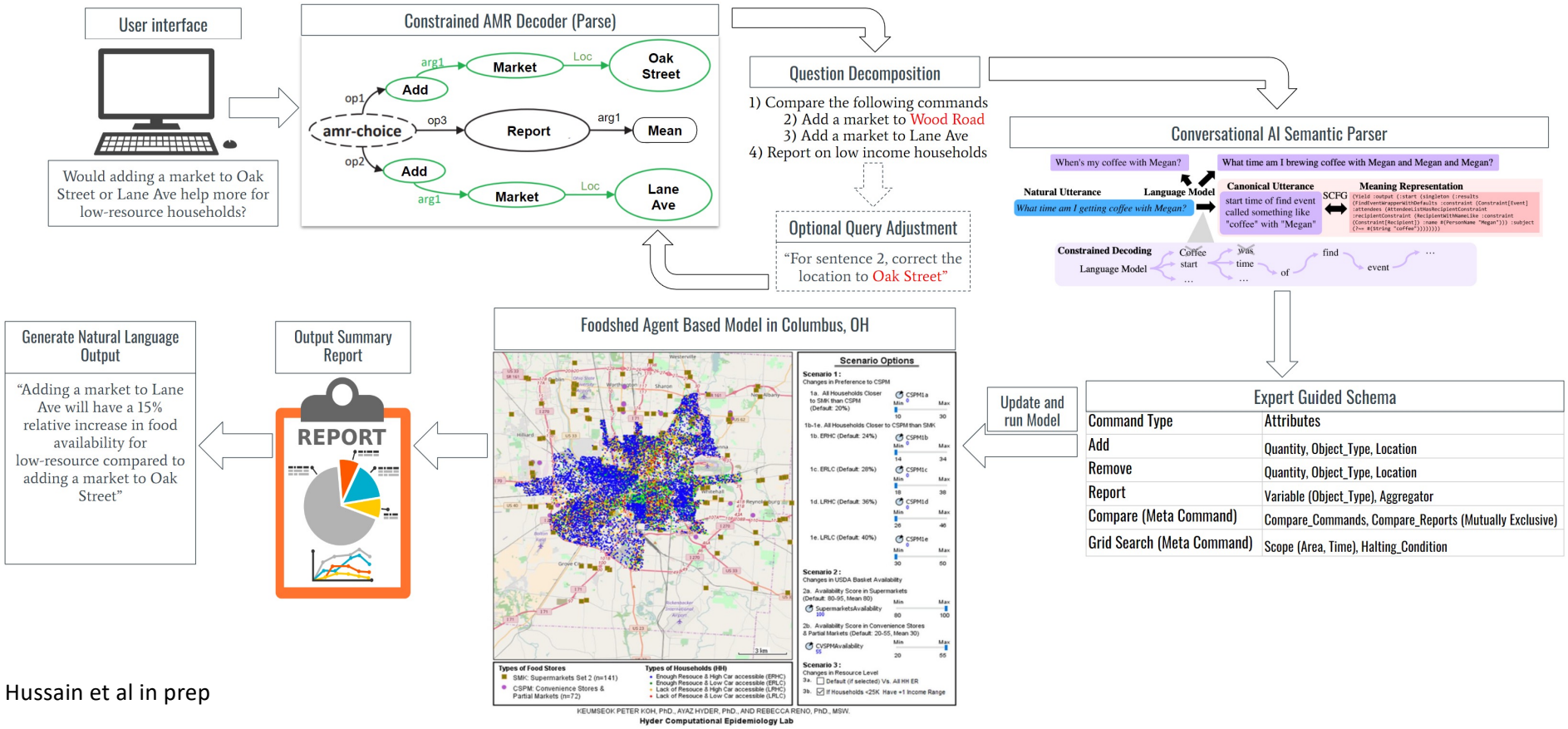
and **long-tail science.**

ICICLE: Intelligent CyberInfrastructure with Computational Learning in the Environment

- "Democratizing AI": how do we bring AI capabilities from HPC center to edge devices to domain scientists working in the field?
  - Digital Agriculture (e.g. drones for crop monitoring)
  - Wildlife Ecology (e.g. camera traps for monitoring wildlife)
  - Smart Foodsheds (e.g. modeling effects of grocery store placements)
- Advancements along many domains needed to make plug-and-play AI accessible to a larger audience
  - Conversational AI plays a role in providing natural interfaces for non-experts
  - Application: Conversational front end for Ayaz's grocery store model
    - How can we get complex questions turned into understandable programs?
    - How can we handle variation in language using pretrained language models (low resource)?
    - Will this model manipulation generalize to other domains?

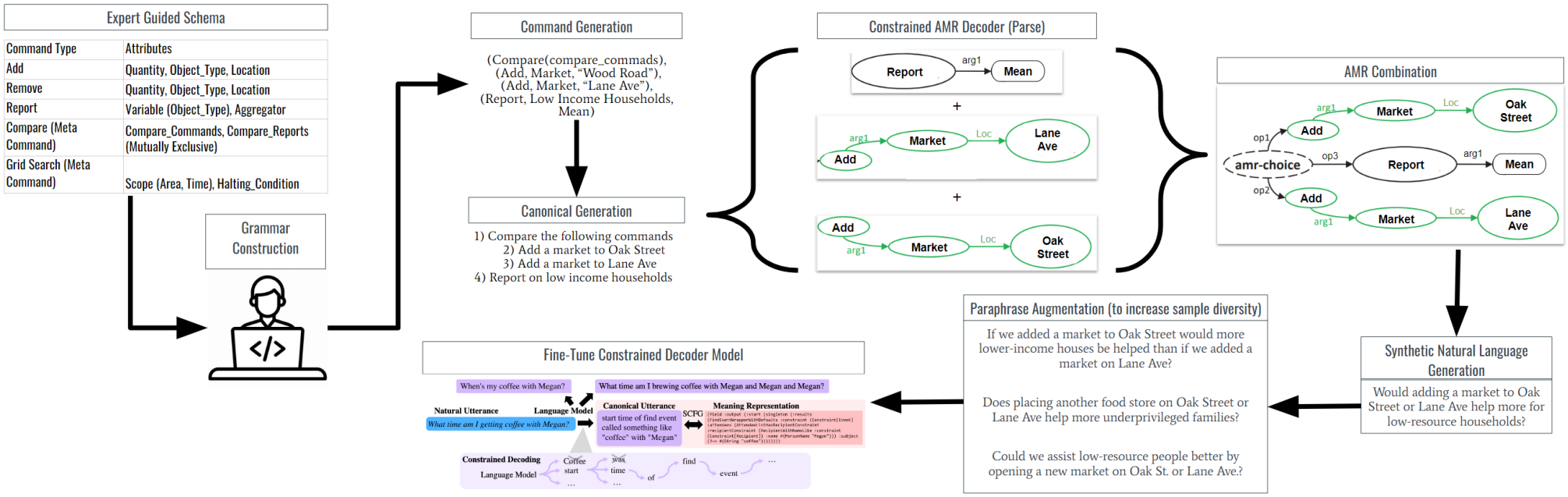
http://icicle.ai

Vision: Getting your model into the hands of others via “Natural Language Programming”



Hussain et al in prep

Vision: Domain experts building their own NLP frontends



Hussain et al in prep

Summary: Interesting challenges in getting NLP into non-NLP researcher's hands

- Partnering with domain experts in medicine, food health has taught us to think differently about building NLP systems

- New research questions: rather than building systems for end users, how do we make it easy for domain experts to develop NLP systems?
  - Transparency and trustworthiness need to be built into the design

- Many domains will provide sparse data for training systems
  - Low resource long tail problems will abound!

Virtual Patient Demo

2020 version: Mr. Rodriguez

Virtual Patient Demo

2020 version: Mr. Rodriguez

Social History

Work/daily activities

No question asked.

Living arrangements and support systems

No question asked.

Physical activity/exercise

No question asked.

Substance use current or past

No question asked.

Diet

No question asked.

Sexuality

No question asked.

Abuse

No question asked.

0 of 7 essential Social History questions asked.

Total questions asked: 4

All Questions Asked:

1. how are you
2. Where does it hurt
3. How long has it been hurting
4. Have you been taking anything for the pain