Masoud Rouhizadeh, Ph.D., MSc, Prof.MS., MA

Malachowsky Hall for Data Science & IT
University of Florida
1889 Museum Rd.
Gainesville, FL 32611

email: mrouhizadeh@ufl.edu
phone: +1 352 273 9397
Website
Google Scholar

RESEARCH INTERESTS

AI in Healthcare, Natural Language Processing, Biomedical Data Science, Social Determinants of Health.

EDUCATION & TRAINING

Postdoctoral Training in Biomedical Informatics & Data Science (2017)
University of Pennsylvania

Ph.D. in Computer Science & Engineering (2015)
Oregon Health & Science University

MSc in Computer Science & Engineering (2013)
Oregon Health & Science University

Professional Master in Human Language Technology (2009)
University of Trento, Italy

MA in Linguistics (2008)
Allameh Tabatabai University

RESEARCH ACTIVITIES

University of Florida

Lead, AI Collaboration Hub, Intelligent Critical Care Center (IC3) 2024–present
Lead, AI Task Force, College of Pharmacy 2023–present
Assistant Professor, Pharmaceutical Outcomes & Policy, UF AI Initiative 2021–present

- Establishing the Florida Pediatric Leukemia Net (Flo_PediaL Net) to investigate biomarkers and socioeconomic factors influencing pediatric leukemia outcomes, utilizing multi-omics analysis.
- Designing a Retrieval-Augmented Generation (RAG) pipeline with knowledge base integration for enhanced extraction of diverse mental health and social determinants from EHR.
- Developing a computable phenotype for drug-induced cardiotoxicity within OneFlorida+ PCORNet, employing advanced NLP/NER for medication and adverse event extraction.
- Collaborating with Harvard to investigate the relationship between social determinants and adolescent mental health outcomes through longitudinal EHR data analysis.
- Leading large language model pipeline development to enhance detection of depression, anxiety, and suicidal behaviors within OneFlorida+ and Mass General Brigham Research Patient Data Registry.
- Collaborating with Harvard to develop scalable NLP pipelines for comprehensive extraction of substance use patterns from EHR data (including frequency, duration, mode of use, and co-use).
- Analyzing unstructured EHR and nursing home data to explore fall circumstances, self-injurious behaviors, and their risk factors in patients with Alzheimer's disease and related dementias (ADRD).
- Investigating individual and health-system factors influencing enrollment and completion rates in the Veterans MOVE! program, with emphasis on disparities and social determinants of health.
- Collaborating with Boston College to identify computational discourse features for targeted language interventions in autism.

Johns Hopkins University

Adjunct Assistant Professor, Biomedical Informatics and Data Science (BIDS) 2021–present

- Course director and instructor, Natural Language Processing in the Health Sciences.
- Directing a multi-site collaborative initiative to extract social determinants of health data from EHRs.

Co-Founder, Johns Hopkins Center for Clinical NLP (C2NLP) 2017–2021
NLP Lead, Institute for Clinical and Translational Research (CTSA)
Faculty Instructor, Biomedical Informatics and Data Science (BIDS)
Affiliations: Center for Language and Speech Processing (CLSP); Center for Population Health IT (CPHIT)

- Built diverse NLP solutions (with spaCy, cTAKES, MetaMap) for clinical text normalization, named entity recognition, and concept linking within the JHU Precision Medicine Analytics Platform.
- Designed an end-to-end NLP pipeline to extract clinical concepts from EHRs, map them to SNOMED-CT, and enable OMOP-compliant data integration for downstream analytics.
- Engineered and deployed large-scale hybrid solutions for de-identifying 350M Epic EHR notes.
- Developed high-throughput pipeline leveraging Docker, Hadoop, HBase - ingested 350M Epic EHR notes, normalized and stored in HL7 CDA (Clinical Document Architecture), preserving section headers, table structures/values, and note metadata.
- Developed EHR analysis models, to analyze disparities in severe hypocalcemia prevalence among minority patients with type 2 diabetes (with FDA CERSIs).
- Led a cross-institutional collaboration (Kaiser Permanente Mid-Atlantic and North California) to develop HIPAA-compliant NLP systems for comprehensive extraction of social determinants of health from multi-site EHR data.
- Developed and evaluated entity linking models using cTAKES and spaCy to detect patient phenotypes, leveraging the JHU PhenoDB and Human Phenotype Ontology (HPO).
- Optimized data processing pipelines for large-scale analytics within PaTH and PCORnet CRNs.
- Contributed to N3C's multi-site NLP framework extracting COVID-19 symptoms, deployed and evaluated on COVID-19 Precision Medicine Analytics Platform Registry (JH-CROWN).
- Developed diverse computational analysis methods to assess language impairments associated with various cognitive conditions such as dementia and mild cognitive impairment.

University of Pennsylvania

Postdoctoral Research Scientist, Penn Institute for Biomedical Informatics 2015–2017
Joint appointment with Stony Brook University from 2015–2016

- Analyzed social media language using dependency parsing, topic modeling, and dimensional reduction methods to identify linguistic predictors of depression in young adults.
- Utilized machine learning techniques to analyze over 2 million social media posts from Twitter and Reddit to identify potential associations between certain medications and congenital heart defects.
- Developed machine learning models to predict locus of control (internal vs. external) from language patterns in social media text, including creation of an annotated dataset.
- Integrated NLP modules into Penn's Differential Language Analysis Toolkit for social media text processing, including components for normalization, segmentation, POS tagging, parsing, relation extraction, and sentiment analysis.

Johns Hopkins University

Senior Affiliate, Center for Language and Speech Processing (CLSP) Summer 2016
Annual Frederick Jelinek Memorial Summer Workshop (JSALT)

- Developed a machine learning model to link social media language features with anxiety symptoms recorded in EHRs, enabling identification of at-risk individuals.

Oregon Health & Science University

Graduate Research Assistant, Computer Science & Engineering 2009-2015

- Developed and evaluated distributional semantic models for characterizing restrictive and repetitive behaviors in autism compared to typical development.
- Developed and evaluated NLP tools to identify conversational reciprocity in spontaneous dialogues from children with autism, aiding assessment of social communication.
- Contributed to the WordsEye text-to-scene system through designing and populating VigNet, a semantic frame-based ontology enabling automated text-to-scene conversion - constructed through information extraction from noisy big data and crowd-sourcing techniques.

Columbia University

Visiting Scholar, Spoken Language Processing Group Summer 2012

- Incorporated VigNet ontology into WordsEye text-to-scene system to enable automated scene visualization from language descriptions.

University of Trento

Graduate Student, Department of Information Engineering & Computer Science 2008-2009
Visiting Researcher, Centre for Mind/Brain sciences (CIMeC)

- Developed a voice-activated restaurant search dialogue system using VoiceXML that enables users to find local restaurants through natural language conversational interactions.
- Developed novel parameterized syntactic tree kernels for question-answer classification, enhancing semantic matching and improving answer retrieval accuracy.
- Improved machine translation performance through text normalization of parallel corpora to reduce morphological ambiguity and improve word alignment.

Fondazione Bruno Kessler (FBK) Trento

Research Intern, Human Language Technology Group Summer 2009

- Developed a high-performance finite-state morphological analyzer for Persian to address the unique challenges of its complex morphology, such as verb compounding constructs.

Shahid Beheshti University

Research Scientist, NLP Group, Computer & Electrical Engineering Department 2007-2009

- Contributed to the development of FarsNet, the first comprehensive Persian WordNet, by applying word sense disambiguation, information retrieval, and machine translation techniques.
- Co-developed the SBUQA system, a question-answering system for open-domain factual questions.

RESEARCH SUPPORT & FUNDING

CURRENT

ID	U18DP006711
Title	*Assessing Barriers and Facilitators for Participating in Structured Lifestyle Intervention and Its Real-world Effectiveness and Cost-effectiveness among US Veterans*
Funding	Centers for Disease Control and Prevention (CDC)
Period	07/2022–06/2027
Role	Co-Investigator (0.10 FTE); PI: Shao H
Total	$2,850,000
Topic	This project aims to identify barriers and facilitators to enrollment and completion of the National Diabetes Prevention Program (NDPP), and to evaluate its long-term effectiveness and cost-effectiveness. The research team will leverage both structured and unstructured data from the Veterans Administration to identify individual and system-level factors associated with enrollment in, and completion of, the NDPP. They will specifically focus on social determinants of health that may be associated with racial/ethnic inequity in program enrollment and completion.

ID	P30AG066506
Title	*1Florida Alzheimer's Disease Research Center (ADRC)*
Funding	NIH National Institute on Aging (NIA)
Period	06/2021–04/2025
Role	Segment PI (FTE reported in the 3 1Florida ADRC projects); PI: Smith
Total	$15,142,000
Topic	This P30 grant supports multiple investigators in enhancing multidisciplinary approaches and collaborative research efforts focused on Alzheimer's Disease and Related Dementias (ADRD).

ID	1Florida ADRC ASCEND
Title	*Advancing Safety in Alzheimer's Disease and Related Disorders – Identifying Fall Circumstances in EHR Data using NLP and Data Science (ASCEND)*
Funding	1Florida Alzheimer's Disease Research Center (ADRC)
Period	05/2024–04/2026
Role	Co-Principal Investigator (0.15 FTE); PIs: Rouhizadeh M, Bjarnadottir R,
Total	$200,000
Topic	Falls are a global public health issue, causing an estimated 684,000 deaths and 38 million disability-adjusted life years lost each year worldwide. Individuals with ADRD are disproportionately affected by falls and fall-related injuries, yet much remains unclear. This study leverages state-of-the-art NLP approaches in EHR data to identify fall events and quantify the circumstances contributing to falls among older adults with ADRD.

ID	1Florida ADRC Suicide
Title	*Identifying Suicidal Behaviors using Natural Language Processing in Patients with Alzheimer's Disease and Related Dementias*
Funding	1Florida Alzheimer's Disease Research Center (ADRC)
Period	01/2022–11/2024
Role	Principal Investigator (0.20 FTE); PI: Rouhizadeh
Total	$200,000
Topic	The overarching goal of this study is to apply, evaluate, and refine the use of a natural language processing-based approach to identify suicide attempts and suicidal ideation in patients with ADRD. The study also plans to examine the epidemiology of suicide attempts and suicidal ideation using the expanded classification and evaluate comparative detection of these behaviors through codified vs. NLP-based approaches.

ID	1Florida ADRC BP
Title	*Identification and Characterization of Blood Pressure Control and Racial Impacts on the Alzheimer's Disease and Related Dementias Risk*
Funding	1Florida Alzheimer's Disease Research Center (ADRC)
Period	01/2022–11/2024
Role	Co-Investigator (0.05 FTE); PI: McDonough C
Total	$200,000
Topic	Hypertension is a potentially modifiable risk factor for ADRD. African Americans and Hispanics have higher rates of ADRD and lower rates of blood pressure control compared to non-Hispanic whites. This project aims to validate and improve ADRD models in these populations, focusing on the role of blood pressure control in risk assessment, to facilitate early identification and targeted treatment.

ID	UF Health GERD
Title	*Can Knowledge of CYP2C19 Genotyping Help Identify Patients with GERD that are Better Managed Surgically?*
Funding	UF Health
Period	11/2023–10/2025
Role	Co-Investigator (0.05 FTE); PI: Puri R
Total	$16,500
Topic	This study aims to: 1) Assess the prevalence of genetic polymorphisms in patients with GERD in the Northern Florida population; 2) Determine the degree of polymorphisms in medical and surgical patients with GERD; 3) Develop a predictive model encompassing anatomic, physiologic, and genetic factors that may be applied clinically to better manage GERD patients. This will represent the first study to compare CYP2C19 polymorphisms in adult surgical versus medical patients with GERD.

PENDING

ID	NIDA R01
Title	*Utilizing Natural Language Processing in Electronic Medical Records to Investigate the Relationship between Cannabis Use and Health Outcomes*
Funding	NIH National Institute on Drug Abuse (NIDA)
Period	07/2024–06/2029
Role	Contact PI (0.30 FTE); Co-PI: Gilman J (Harvard)
Total	$4,124,000
Topic	Cannabis use has increased in prevalence across most states in the United States since recreational or medicinal use was legalized in all but eight states by 2023. This proposal aims to investigate the impacts of cannabis use on mental health outcomes among young adults, examining its relationships with new-onset psychosis, self-harm, and substance use disorders, while also assessing how social determinants like income and housing modify these associations. Findings would provide real-world evidence to inform clinical practice, population health studies, and policy regarding rising cannabis use and associated mental health risks in this demographic.

ID	NIA R01
Title	*Harnessing Real-World Data for Enhanced Frailty Assessment and Surgical Outcomes*
Funding	NIH National Institute on Aging (NIA)
Period	01/2025–12/2029
Role	Co-Investigator (0.25 FTE); PI: Mardini M. (UF) -
Total	$3,361,900
Topic	This project focuses on improving preoperative care for older surgical patients by developing an automated machine-learning tool for assessing frailty, a vital health indicator that affects surgical outcomes. By integrating real-world data and advanced analytics, this tool aims to make frailty evaluations more accurate and practical, ultimately leading to personalized treatment plans and better post-surgery outcomes. This research is significant for public health as it seeks to enhance healthcare efficiency, reduce postoperative complications, and improve the quality of life for older adults undergoing surgery.

ID	FL DOM Flo_PediaL Net
Title	*Florida Pediatric Leukemia Net (Flo_PediaL Net)*
Funding	Florida Department of Health
Period	5/2024–4/2025
Role	Co-Investigator (0.15 FTE); PI: Lamba J
Total	$224,300
Topic	The proposed project aims to define biomarkers and socioeconomic factors impacting outcomes in pediatric acute myeloid leukemia (AML) cases in Florida. Integrated multi-omics analysis of genomic, transcriptomic, proteomic, and metabolomic data from pediatric AML patients will establish predictive markers of prognosis, treatment response, and novel therapeutic targets. Additionally, parameters in the OneFlorida dataset will be analyzed to determine structural factors associated with disparities in pediatric leukemia outcomes across Florida. Results on omics-based biomarkers and non-biological factors impacting survival rates will ultimately inform efforts to improve risk-stratification and reduce outcome inequities for pediatric AML patients in Florida.

ID	DoD-JHU
Title	*Assessing the Impact of Symptom Burden and Treatment on Disability and Progression in Multiple Sclerosis*
Funding	U.S. Department of Defense (DoD) - Johns Hopkins University
Period	09/2024–09/2026
Role	Principal Investigator (0.10 FTE)
Total	$280,000
Topic	We will develop large language models to identify hidden symptoms of multiple sclerosis from free text clinical notes within electronic health records. We will build and evaluate large language models on Johns Hopkins MS Center visit notes to expand data collection of fatigue, depression, cognitive dysfunction, pain, anxiety, spasticity, and bowel/bladder dysfunction. This will enable us to conduct a case-control study assessing the impact of symptom burden and treatment on disability and disease progression in MS.

ID	DoD-UF
Title	*Leveraging Large Language Models for Summarizing Electronic Clinical Note*
Funding	U.S. Department of Defense (DoD)
Period	06/2024–09/2024
Role	Co-Investigator; (0.05 FTE) PI: Mardini M
Total	$100,000
Topic	Healthcare professionals face overwhelming data volumes in electronic records, hampering extraction of essential details. We propose automatically creating summaries of complex clinical narratives via Large Language Models, improving information retrieval and decisions. Summaries will be iteratively refined through physician-in-the-loop evaluation assessing conciseness, relevance, readability, and incorporation of expert feedback. This research can offer insights on using language models for clinical summarization, develop physician-validated standards for implementation, and inform best practices for integrating these tools into healthcare systems.

ID	NRSA Fellowship
Title	*Identifying and Predicting Bleeding Risk with Prasugrel and Ticagrelor after Percutaneous Coronary Intervention*
Funding	NIH National Research Service Award (NRSA)
Period	09/2024–08/2026
Role	Advisory Committee Member (pre-doc) (in-kind FTE); PI: Cameron T
Total	$106,600
Topic	There are important gaps in the evidence on patient-specific factors influencing the safety of antiplatelet therapy with prasugrel and ticagrelor after percutaneous coronary intervention. This proposal will identify clinical factors influencing bleeding risk with prasugrel or ticagrelor from large-scale health record data and genetic associations with bleeding in a diverse real-world population. Through completion of this work, I will obtain the skills and experience needed to identify genetic and other patient-specific contributors to the inter-patient variability in drug response.

COMPLETED

ID	1R01DA050676
Title	*Developing and Evaluating a Machine-Learning Opioid Prediction & Risk-Stratification E-Platform (DEMONSTRATE)*
Funding	NIH National Institute on Drug Abuse (NIDA)
Period	07/2021–10/2023
Role	Co-Investigator; PI: Lo-Ciganic J
Total	$3,260,000
Topic	The proposed study aimed to harness advanced natural language processing and longitudinal neural network approaches to build on our previously developed machine learning prediction algorithms for identifying patients at risk for opioid overdose or opioid use disorder. We developed the prediction tool using all-payer EHRs, Medicaid claims, and Medicaid claims linked with EHR data from the One Florida Clinical Research Consortium and translated the risk prediction algorithms into a clinical decision support platform integrated into the EHR system to identify patients at high risk of overdose and opioid use disorder.

ID	1R01MD015844
Title	*Development, Piloting, and Dissemination of an Integrated Clinical and Social Multi-Level Decision Support Platform to Address Social Determinants of Health among Minority Populations in Baltimore City*
Funding	NIH National Institute of Mental Health (NIMH)
Period	07/2021—03/2023
Role	Co-Investigator; PIs: Weiner J, Hatef E (JHU)
Total	$2,746,000
Topic	With the intent of improving care for minority and disadvantaged populations with chronic diseases, this project developed methods and tools to better integrate available digital information regarding social determinants of health (SDOH) into providers' electronic health records (EHRs). The study developed, piloted, evaluated, and widely disseminated a multi-level EHR-integrated clinical decision support system to help identify, manage, and refer patients with a high chronic disease burden who also had high levels of modifiable SDOH challenges.

ID	5U01FD005942-05
Title	*Assessing Disparities in Occurrence and Outcomes of Type 2 Diabetes Adverse Events in Minority Populations Using Real World Administrative Claims and Electronic Health Records*
Funding	FDA Centers of Excellence in Regulatory Science and Innovation (CERSIs)
Period	09/2020—09/2021
Role	Co-Investigator; PI: Kharrazi H, Weiner J (JHU)
Total	$741,000
Topic	This project aimed to improve the identification of severe hypoglycemia (SHG) in ambulatory EHRs and claims. We compared SHG events across races/ethnicities and various social determinants of health (SDH) factors, and then identified key disparity factors associated with an increased likelihood of SHG among African American patients. We also discovered contextual patterns associated with higher rates of SHG among different minority and special-needs populations.

ID	R01MH124724
Title	*Advancing Maryland's Statewide Suicide Data Warehouse to Improve Individual and Population-Level Mortality Prediction and Prevention*
Funding	NIH National Institute of Mental Health (NIMH)
Period	09/2020—09/2021
Role	Co-Investigator; PI: Kharrazi
Total	$3,260,000
Topic	The aim of developing the Maryland Suicide Data Warehouse (MSDW) was to link a diverse set of data sources and use various layers of risk factors to predict suicide mortality on a generalizable population. MSDW included data on approximately 5+ million Maryland residents spanning from 2012 to 2017. MSDW contained hospital and emergency discharges, health information exchange (HIE) data, commercial insurance claims, medical examiner data, EHR data from 5 select health systems, and place-based social determinants of health data covering various domains such as housing, employment, education, income and crime. This study leveraged the unique data types linked by MSDW to improve predictions of suicide deaths among generalizable patient populations.

ID	HSRP20202805
Title	*Using Topic Segmentation to Improve Concept Parsing and Identification of Negation in Extraction of Data from EHRs*
Funding	Patient-Centered Outcomes Research Institute (PCORI)
Period	11/2019—09/2021
Role	Project Leader; PIs: Turchin A (Harvard), Ford D (JHU)
Total	$1,263,012
Topic	A major challenge when using EHR data for research was that a large fraction of the data was contained in free-text documents. Even with modern NLP methods, it could be difficult to extract information when topics switched mid-sentence. The goal of this project was to build NLP for topic segmentation to help with better clinical concept extraction.

ID	1R56-MH117560-01
Title	*Addressing Suicide Research Gaps: Understanding Mortality Outcomes in the Mid-Atlantic Region*
Funding	NIH National Institute of Mental Health (NIHM)
Period	09/2018–09/2021
Role	Co-Investigator; PIs: Kharrazi H, Wilcox H
Total	$492,000 (R56 mechanism – FY1 of 1R01-MH117560-01)
Topic	This study conducted data linkage and informatics approaches to utilize existing resources to improve suicide risk identification and prevention. The project also offered essential measures for providers and payers to help reduce suicide events in their systems. The methods were assessed against and shared as a resource with other states/regions to establish a similar analytics framework to advance suicide prevention efforts.

ID	UFII Suicide
Title	*Diagnosing Suicidal Behaviors in Postpartum Mothers Using Natural Language Processing*
Funding	Internal - UF Informatics Institute (UFII)
Period	01/2022–01/2023
Role	Principal Investigator; Co-PI: Brown J
Total	$28,000
Topic	In this project, we refined an existing NLP text mining and heuristic rule-based algorithm that identified suicide attempts and behaviors in postpartum mothers, and applied the refined NLP tool to evaluate demographic and clinical characteristics associated with suicide attempts and suicide-related behaviors.

ID	UFII P2Y12
Title	*Utilizing Artificial Intelligence to Identify Bleeding Risk Predictors with Newer P2Y12 Receptor Inhibitors*
Funding	Internal - UF Informatics Institute (UFII)
Period	01/2022–01/2023
Role	Co-Investigator; Co-PIs: McDonough C, Cavallari L
Total	$25,000
Topic	This project aimed to identify patient-specific factors predictive of bleeding risk with prasugrel or ticagrelor after PCI through AI approaches; and to derive a novel score for predicting bleeding risk with prasugrel and ticagrelor that could be integrated into the EHR to provide point-of-care risk assessment.

ID	JHU Kaiser Permanente
Title	*Applying Unstructured EHR Notes and Structured Clinical and Community Data to Identify Patient Social Needs and Determinants*
Funding	Internal - Johns Hopkins ICTR & Kaiser Permanente Mid-Atlantic
Period	04/2020–04/2021
Role	Co-Investigator; PIs: Hatef E, Nau C
Total	$75,000
Topic	The overarching goal of this project was to develop and apply new practical tools and approaches for measuring Social Determinants of Health using unstructured (e.g., clinician notes) and structured EHR data and information derived from neighborhood databases.

ID	JHSoM ICTR BSSS Nexus Award [Phase III]
Title	*Behavioral, Social, and Systems Science: Extracting Social Science Data from Epic Electronic Health Record System*
Funding	Internal - Johns Hopkins Institute for Clinical and Translational Research (ICTR)
Period	04/2020–04/2021
Role	Co-Investigator; PI: Kharrazi H
Total	$95,000
Topic	This phase of the project completed extraction of an array of social and behavioral determinants data from the EPIC EHR, spanning factors like financial strain, housing issues, and transport barriers. Additionally, we integrated external location-based metrics, capturing community-level influences on health across the system's area.

ID	JHU PMAP JHCROWN
Title	*AI and NLP Methods for Extracting COVID-19 Signs and Symptoms from Free Text Clinical Notes*
Funding	Internal - Johns Hopkins Precision Medicine Analytics Platform (PMAP)
Period	04/2020–03/2021
Role	Project Lead
Total	$36,000
Topic	Statistical modeling of patient trajectories among persons hospitalized for COVID-19 was critical to understand disease progression and provide better clinical care. JHU clinical and biostatistics experts created a model for COVID-19 progression to severe disease or death. Factors included demographics, vitals, comorbid conditions as well as presenting symptoms extracted from hospital notes using NLP methods.

ID	JHU PMAP UMLS
Title	*Semantic Tagging and Concept Extraction from Free-Text Clinical Notes in the JMHI Health Systems*
Funding	Internal - Johns Hopkins Precision Medicine Analytics Platform (PMAP)
Period	03/2019–02/2021
Role	Project Lead
Total	$25,000
Topic	The JHMI EHR system already contained 350 million clinical notes and on average 70,000 notes were added daily. Our enterprise-level big-data concept extraction tool was designed to extract clinical concepts and convert imperceptible unstructured data into a structured format, linked to the NIH UMLS standardized vocabulary.

ID	JHU PMAP HL7 CDA
Title	*Standardizing Clinical Notes Within HL7 Clinical Document Architecture (CDA)*
Funding	Internal - Johns Hopkins Precision Medicine Analytics Platform (PMAP)
Period	08/2019–07/2020
Role	Project Lead
Total	$30,000
Topic	The goal of this project was to build a parser to convert Epic-generated Rich Text Format (RTF) notes into semi-structured HL7 Clinical Document Architecture (CDA) format, and preserve valuable information stored in RTFs such as section headers, section boundaries, and tabular formats. We applied this process to the enterprise-level repository of all 400 million clinic notes in JHMI.

ID	JHU IDIES
Title	*Harnessing Big Data for Population Health: Advancing Natural Language Processing Techniques to Extract Social-Behavioral Risk Factors from Free Text within Large Electronic Health Record Systems*
Funding	Internal - Johns Hopkins Institute for Data-Intensive Engineering and Science (IDIES)
Period	04/2018–03/2019
Role	Co-Investigator; PI: Weiner J
Total	$25,000
Topic	We built on the machine learning tools developed using a mix of pattern matching techniques for note categorization and deep learning methods. Our goal in this pilot was to establish the viability of these methods for extracting social determinants of health (SDOH) data and to extract a first-pass corpus of SDOH factors from clinical notes.

ID	CCDA Core Coins
Title	*Center for Clinical Data Analysis (CCDA) Core Coins*
Funding	Internal - Johns Hopkins Institute for Clinical and Translational Research (ICTR)
Period	01/2018–12/2019
Role	Project Lead
Total	$50,000 (distributed into five projects at $10,000 each)
Topic	Our goal in this program was to provide early-career researchers the opportunity to utilize NLP tools and methodologies to analyze clinical notes to support their research. This initiative aimed to bridge the gap between cutting-edge technology and practical healthcare applications, enabling these researchers to uncover new insights from unstructured clinical data.

ID	JHSoM ICTR BSSS Nexus Award [Phase II]
Title	*Behavioral, Social, and Systems Science: Extracting Social Science Data from Epic Electronic Health Record System*
Funding	Internal - Johns Hopkins Institute for Clinical and Translational Research (ICTR)
Period	12/2017–09/2018
Role	Co-Investigator; PI: Kharrazi, H
Total	$65,000
Topic	In this phase of the project, we documented the array of social determinant measures currently available across various datasets, spanning access, financial, housing, transportation, and other key domains. We provided researchers a practical guide to integrating existing measures into studies on outcomes of interest. Additionally, we put forward focused recommendations on expanding capture of salient social determinants going forward—like food security, neighborhood factors, social connection, and structural inequities. Enhancing our social determinants data assets will strengthen future analyses seeking to explain health disparities.

ID	TRT0048
Title	*Measuring Well-Being Using Big Data, Social Media, and Language Analyses*
Funding	Templeton Religion Trust
Period	11/2015–02/2017
Role	Postdoctoral Research Associate; PI: Seligman, M
Total	$3,800,000
Topic	Founding the World Well-Being Project at the University of Pennsylvania, we developed new techniques for measuring psychological and medical well-being based on language in social media. We investigated how psychosocial processes affected health and happiness, and developed unobtrusive well-being measures to supplement expensive survey methods.

PUBLICATIONS

BOOK CHAPTERS

1. **Rouhizadeh M**, Bowler M, Sproat R, Coyne B. *Data collection and normalization for building the scenario-based lexical knowledge resource of a text-to-scene conversion system.* In A. König et al. (Eds.), Semantic Media Adaptation and Personalization (pp. 147-152). Springer, Berlin, Heidelberg; 2011.
DOI: 10.1109/SMAP.2010.5706851

2. **Rouhizadeh M**, Coyne B, Sproat R. *Collecting semantic information for locations in the scenario-based lexical knowledge resource of a text-to-scene conversion system.* In A. König et al. (Eds.), Lecture Notes in Artificial Intelligence (Vol. 6884, pp. 368-379). Springer-Verlag, Berlin, Heidelberg; 2011.
DOI: 10.1007/978-3-642-23866-6_40

3. Yarmohammadi M, Shamsfard M, Yarmohammadi M, **Rouhizadeh M**. *SBUQA question answering system.* In H. Sarbazi-Azad et al. (Eds.), Advances in Computer Science and Engineering, Communications in Computer and Information Science, (CCIS Vol. 6, pp. 369-377). Springer, Berlin, Heidelberg; 2009.
DOI: 10.1007/978-3-540-89985-3_39

JOURNAL PAPERS

1. Lemas DJ, Du X, **Rouhizadeh, M**, Lewis B, Frank S, Wright L, Spirache A, Gonzalez L, Cheves R, Magalhães M, Zapata R, Reddy R, Xu K, Parker L, Harle C, Young B, Louis-Jaques A, Zhang B, Thompson LA, Hogan WR, Modave F. *Classifying early infant feeding status from clinical notes using natural language processing and machine learning.* Scientific Reports; 2024.
PMID: Pending.

2. Gray GM, Zirikly A, Ahumada LM, **Rouhizadeh M**, Richards T, Kitchen C, Foroughmand I, Hatef E. *Application of natural language processing to identify social needs from patient medical notes: development and assessment of a scalable, performant, and rule-based model in an integrated healthcare delivery system.* JAMIA Open. 2023; 6(4):ooad085.
PMID: 37799347 DOI: 10.1093/jamiaopen/ooad085

3. Liu S, Wen A, Wang L, He H, Fu S, Miller R, Williams A, Harris D, Kavuluru R, Liu M, Abu-El-Rub N, Schutte D, Zhang R, **Rouhizadeh M**, Osborne JD, He Y, Topaloglu U, Hong SS, Saltz JH, Schaffter T, Pfaff E, Chute CG, Duong T, Haendel MA, Fuentes R, Szolovits P, Xu H, Liu H. *An open natural language processing (NLP) framework for EHR-based clinical research: a case demonstration using the*

National COVID Cohort Collaborative (N3C). J Am Med Inform Assoc. 2023; 30(12):2036-2040.
PMID: 37555837 DOI: 10.1093/jamia/ocad134

4. Omaki E, Shields W, **Rouhizadeh M**, Delgado-Barroso P, Stefanos R, Gielen A. *Understanding the circumstances of paediatric fall injuries: a machine learning analysis of NEISS narratives*. Inj Prev. 2023; 29(5):384-388.
PMID: 37399309 DOI: 10.1136/ip-2023-044858

5. Zolnour A, Eldredge CE, Faiola A, Yaghoobzadeh Y, Khani M, Foy D, Topaz M, Kharrazi H, Fung KW, Fontelo P, Davoudi A, Tabaie A, Breitinger SA, Oesterle TS, **Rouhizadeh M**, Zonnor Z, Moen H, Patrick TB, Zolnoori M. *A risk identification model for detection of patients at risk of antidepressant discontinuation*. Front Artif Intell. 2023; 6(1229609).
PMID: 37693012 DOI: 10.3389/frai.2023.1229609

6. Sajdeya R, Mardini MT, Tighe PJ, Ison RL, Bai C, Jugl S, Hanzhi G, Zandbiglari K, Adiba FI, Winterstein AG, Pearson TA, Cook RL, **Rouhizadeh M**. *Developing and validating a natural language processing algorithm to extract preoperative cannabis use status documentation from unstructured narrative clinical notes*. J Am Med Inform Assoc. 2023; 30(8):1418-1428.
PMID: 37178155 DOI: 10.1093/jamia/ocad080

7. Keloth VK, Banda JM, Gurley M, Heider PM, Kennedy G, Liu H, Liu F, Miller T, Natarajan K, V Patterson O, Peng Y, Raja K, Reeves RM, **Rouhizadeh M**, Shi J, Wang X, Wang Y, Wei WQ, Williams AE, Zhang R, Belenkaya R, Reich C, Blacketer C, Ryan P, Hripcsak G, Elhadad N, Xu H. *Representing and utilizing clinical textual data for real world studies: An OHDSI approach*. J Biomed Inform. 2023; 142:104343.
PMID: 36935011 DOI: 10.1016/j.jbi.2023.104343

8. Hatef E, **Rouhizadeh M**, Nau C, Xie F, Rouillard C, Abu-Nasser M, Padilla A, Lyons LJ, Kharrazi H, Weiner JP, Roblin D. *Development and assessment of a natural language processing model to identify residential instability in electronic health records' unstructured data: a comparison of 3 integrated healthcare delivery systems*. JAMIA Open. 2022; 5(1):ooac006.
PMID: 35224458 DOI: 10.1093/jamiaopen/ooac006

9. Schwartz JL, Tseng E, Maruthur NM, **Rouhizadeh M**. *Identification of prediabetes discussions in unstructured clinical documentation: validation of a natural language processing algorithm*. JMIR Med Inform. 2022; 10(2):e29803.
PMID: 35200154 DOI: 10.2196/29803

10. Alipour-Haris G, Armstrong MJ, Sullivan JL, Suryadevara U, **Rouhizadeh M**, Brown JD. *Suicidal ideation and suicide-attempt-related hospitalizations among people with Alzheimer's disease (AD) and AD-related dementias in the United States during 2016-2018*. J Clin Med. 2022; 11(4):943.
PMID: 35207214 DOI: 10.3390/jcm11040943

11. Hatef E, **Rouhizadeh M**, Nau C, Xie F, Padilla A, Lyons LJ, Rouillard C, Abu-Nasser M, Roblin D. *A pilot study to improve the use of electronic health records for identification of patients with social determinants of health challenges: a collaboration of Johns Hopkins Health System and Kaiser Permanente*. Health Serv Res. 2021 Sep;56(Suppl 2):27–8.
PMCID: PMC8441436 DOI: 10.1111/1475-6773.13756

12. Hatef E, Singh Deol G, **Rouhizadeh M**, Li A, Eibensteiner K, Monsen CB, Bratslaver R, Senese M, Kharrazi H. *Measuring the value of a practical text mining approach to identify patients with housing issues in the free-text notes in electronic health record: findings of a retrospective cohort study*. Front Public Health. 2021; 9:697501.
PMID: 34513783 DOI: 10.3389/fpubh.2021.697501

13. Wang J, Abu-El-Rub N, Gray J, Pham HA, Zhou Y, Manion FJ, Liu M, Song X, Xu H, Zhang Y, **Rouhizadeh M** *COVID-19 SignSym: a fast adaptation of a general clinical NLP tool to identify and normalize COVID-19 signs and symptoms to OMOP common data model.* J Am Med Inform Assoc. 2021; 28(6):1275-1283.
PMID: 33674830 DOI: 10.1093/jamia/ocab015

14. Hatef E, Ma X, **Rouhizadeh M**, Singh G, Weiner JP, Kharrazi H. *Assessing the impact of social needs and social determinants of health on health care utilization: using patient- and community-level data.* Popul Health Manag. 2021; 24(2):222-230.
PMID: 32598228 DOI: 10.1089/pop.2020.0043

15. Tseng E, Schwartz JL, **Rouhizadeh M**, Maruthur NM. *Analysis of primary care provider electronic health record notes for discussions of prediabetes using natural language processing methods.* J Gen Intern Med. 2021; DOI: 10.1007/s11606-020-06400-1.
PMID: 33469758 DOI: 10.1007/s11606-020-06400-1

16. Garibaldi BT, Fiksel J, Muschelli J, Robinson ML, **Rouhizadeh M**, Perin J, Schumock G, Nagy P, Gray JH, Malapati H, Ghobadi-Krueger M, Niessen TM, Kim BS, Hill PM, Ahmed MS, Dobkin ED, Blanding R, Abele J, Woods B, Harkness K, Thiemann DR, Bowring MG, Shah AB, Wang MC, Bandeen-Roche K, Rosen A, Zeger SL, Gupta A. *Patient trajectories among persons hospitalized for COVID-19: a cohort study.* Ann Intern Med. 2021; 174(1):33-41.
PMID: 32960645 DOI: 10.7326/M20-3905

17. Dickerson LK, **Rouhizadeh M**, Korotkaya Y, Bowring MG, Massie AB, McAdams-Demarco MA, Segev DL, Cannon A, Guerrerio AL, Chen PH, Philosophe BN, Mogul DB. *Language impairment in adults with end-stage liver disease: application of natural language processing towards patient-generated health records.* NPJ Digit Med. 2019; 2:106.
PMID: 31701020 DOI: 10.1038/s41746-019-0179-9

18. Hatef E, **Rouhizadeh M**, Tia I, Lasser E, Hill-Briggs F, Marsteller J, Kharrazi H. *Assessing the availability of data on social and behavioral determinants in structured and unstructured electronic health records: a retrospective analysis of a multilevel health care system.* JMIR Med Inform. 2019; 7(3):e13802.
PMID: 31376277 DOI: 10.2196/13802

19. Mansoory N, Shamsfard M, **Rouhizadeh M**. *Compound verbs in Persian Wordnet.* International Journal of Lexicography, Oxford University Press. 2012 Mar;25(1):50-67.
DOI: 10.1093/ijl/ecr022

TOP-TIER CONFERENCE PAPERS

My publications in highly selective top-tier computer science conferences demonstrate significant impact. These venues hold high h5-indices in Google Scholar, often comparable to or exceeding top biomedical journals such as *Nature Biotechnology* (h5-index: 185) and the *European Heart Journal* (h5-index: 171). With acceptance rates around 20%, these prestigious conferences are renowned for their rigorous peer-review processes, on par with the selectivity of leading journals.

1. Mulyar A, Schumacher E, **Rouhizadeh M**, Dredze M. *Phenotyping of clinical notes with improved document classification models using contextualized neural language models.* Neural Information Processing Systems (NeurIPS). 2019.
DOI: 10.48550/arXiv.1910.13664
CORE Rank: A* Acceptance Rate: 21% h5-index: 309

2. **Rouhizadeh M**, Jaidka K, Smith L, Schwartz HA, Buffone A, Ungar L. *Identifying locus of control in social media language.* Empirical Methods in Natural Language Processing (EMNLP). 2018.
DOI: 10.18653/v1/D18-1145
CORE Rank: A* Acceptance Rate: 23% h5-index: 176

3. Jaidka K, Buffone A, Eichstaedt J, **Rouhizadeh M**, Ungar L. *Modeling and visualizing locus of control with Facebook language.* Association for the Advancement of Artificial Intelligence (AAAI). 2018.
DOI: 10.1609/icwsm.v12i1.15076
CORE Rank: A* Acceptance Rate: 20% h5-index: 212

4. Schwartz HA, **Rouhizadeh M**, Bishop M, Tetlock P, Mellers B, Ungar L. *Assessing objective recommendation quality through political forecasting.* Empirical Methods in Natural Language Processing (EMNLP). 2017.
DOI: 10.18653/v1/D17-1250
CORE Rank: A* Acceptance Rate: 18% h5-index: 176

5. **Rouhizadeh M**, Ungar L, Buffone A, Schwartz HA. *Using syntactic and semantic context to explore psychodemographic differences in self-reference.* Empirical Methods in Natural Language Processing (EMNLP). 2016.
DOI: 10.18653/v1/D16-1219
CORE Rank: A* Acceptance Rate: 22% h5-index: 176

6. **Rouhizadeh M**, Prud'hommeaux E, van Santen J, Sproat R. *Measuring idiosyncratic interests in children with autism.* Association for Computational Linguistics (ACL). 2015 Jul;2015:212-217.
PMID: 29217874 DOI: 10.3115/v1/p15-2035
CORE Rank: A* Acceptance Rate: 22% h5-index: 192

7. **Rouhizadeh M**, Sproat R, van Santen J. *Similarity measures for quantifying restrictive and repetitive behavior in conversations of autistic children.* North American Chapter of the Association for Computational Linguistics (NAACL-HLT). 2015 Jun 5;2015:117-123.
PMID: 28691123 DOI: 10.3115/v1/N15-1013
CORE Rank: A Acceptance Rate: 22% h5-index: 133

8. **Rouhizadeh M**, Prud'hommeaux E, Roark B, van Santen J. *Distributional semantic models for the evaluation of disordered language.* North American Chapter of the Association for Computational Linguistics (NAACL-HLT). 2013 Jun;2013:709-714.
DOI: 10.3115/v1/N13-1087 PMID: 25419547
CORE Rank: A Acceptance Rate: 22% h5-index: 133

9. Coyne B, Klapheke A, **Rouhizadeh M**, Sproat R, Bauer D. *Annotation tools and knowledge representation for a text-to-scene system.* Computational Linguistics (COLING). 2012.
ACL: C12-1042
CORE Rank: B Acceptance Rate: 24% h5-index: 73

ADDITIONAL REFEREED FULL CONFERENCE PAPERS

1. Hossein R, Shamsfard M, **Rouhizadeh M**. *Persian SemCor: a bag of word sense annotated corpus for the Persian language.* Global WordNet Conference (GWC). 2021.
ACL: 2021.gwc-1.17

2. Hossein R, Shamsfard M, **Rouhizadeh M**. *Knowledge based word sense disambiguation with distributional semantic expansion for the Persian language.* International eConference on Computer and

Knowledge Engineering (ICCKE). 2020.
URL: ICCKE-2020

3. Hossein R, Shamsfard M, **Rouhizadeh M**. *Knowledge-based word sense disambiguation with distributional semantic expansion.* Workshop on Widening NLP in Association for Computational Linguistics (WiNLP). 2019.
ACL: W19-3604

4. **Rouhizadeh M**, Magge A, Klein A, Sarker A, Gonzalez G. *A rule-based approach to determining pregnancy time-frame from contextual social media postings.* ACM International Digital Health. 2018.
ACM: DH-2018

5. Klein AZ, Sarker A, **Rouhizadeh M**, O'Connor K, Gonzalez G. *Detecting personal medication intake in Twitter: an annotated corpus and baseline classification system.* Biomedical Natural Language Processing (BioNLP). 2017.
DOI: 10.18653/v1/W17-2316

6. Prud'hommeaux E, Morley E, **Rouhizadeh M**, Silverman L, van Santen J, Roark B, Sproat R, Kauper S, DeLaHunta R. *Computational analysis of trajectories of linguistic development in autism.* IEEE Spoken Language Technology (SLT). 2014 Dec;2014:266-271.
PMID: 29057398 DOI: 10.1109/SLT.2014.7078585

7. **Rouhizadeh M**, Prud'hommeaux E, van Santen J, Sproat R. *Detecting linguistic restricted interests in autism using distributional semantic models.* Computational Linguistics and Clinical Psychology (CLPsych). 2014.
DOI: 10.3115/v1/W14-3206

8. Prud'hommeaux E, **Rouhizadeh M**. *Automatic detection of pragmatic deficits in children with autism.* Workshop on Child Computer Interaction (WOCCI). 2012 Sep 14;2012:1-6.
PMID: 28691126 PMCID: PMC5500165

9. **Rouhizadeh M**, Bauer D, Coyne B, Rambow O, Sproat R. *Collecting spatial information for locations in a text-to-scene conversion system.* Workshop on Computational Models of Spatial Language Interpretation and Generation (CoSLI). 2011.
URL: CoSLI-2011

10. **Rouhizadeh M**, Bowler M, Sproat R, Coyne B. *Collecting semantic data from Amazon Mechanical Turk for a lexical knowledge resource in a text to picture generating system.* International Conference on Computational Semantics (IWCS). 2011.
ACL: W11-0147

11. **Rouhizadeh M**, Yarmohammadi M, Shamsfard M. *Developing the Persian WordNet of verbs.* Global WordNet Association. 2010.
URL: GWC-2010

12. Yarmohammadi M, Shamsfard M, **Rouhizadeh M**. *Using WordNet in extracting the final answer from retrieved documents in a question answering system.* Global WordNet Association. 2008.
URL: GWC-2008-QA

13. **Rouhizadeh M**, Shamsfard M, Yarmohammadi M. *Building a WordNet for Persian verbs.* Global WordNet Association. 2008.
URL: GWC-2008-WN

14. **Rouhizadeh M**, Assi M, Yarmohamadi MA. *Designing Persian verbs WordNet.* Iranian Conference

on Linguistics. 2007.
URL: ICL-2007-PWN

PEER-REVIEWED ABSTRACTS

1. Zandbiglari K, **Rouhizadeh M**. *Enhancing suicide risk detection in EHRs: leveraging large language models to identify explicit and implicit suicidal indications.* American Medical Informatics Association Informatics Summit. 2024.

2. Zandbiglari K, **Rouhizadeh M**. *Leveraging large language models to identify suicidal behaviors from electronic health records.* UF College of Pharmacy Annual Research Showcase. 2024.

3. Adiba F, **Rouhizadeh M**. *Large language models for identifying social determinants of health in postpartum depression.* UF College of Pharmacy Annual Research Showcase. 2024.

4. Sajdeya R, Bai C, Jugl S, Ison RL, Mardini MT, Tighe PJ, Zandbiglari K, Gao H, Winterstein AG, Pearson TA, Cook RL, **Rouhizadeh M**. *Preoperative cannabis use status ascertainment using natural language processing methods and coded data.* American Medical Informatics Association Annual Symposium (AMIA). 2023. ***Best Poster Award Nominee***

5. Zandbiglari K, Omaki E, Adiba FI, Shields WC, Gielen A, **Rouhizadeh M**. *Deep learning and natural language processing methods for characterizing pediatric falls events.* American Medical Informatics Association Annual Symposium (AMIA). 2023.

6. Sajdeya R, Bai C, Jugl S, Ison RL, Mardini MT, Tighe PJ, Zandbiglari K, Gao H, Winterstein AG, Pearson TA, Cook RL, **Rouhizadeh M**. *Development and validation of a natural language processing algorithm for extracting preoperative cannabis use status information from unstructured narrative clinical notes.* American Medical Informatics Association Annual Symposium (AMIA). 2023.

7. Sajdeya R, Bai C, Jugl S, Ison RL, Mardini MT, Tighe PJ, Zandbiglari K, Gao H, Winterstein AG, Pearson TA, Cook RL, **Rouhizadeh M**. *Developing and validating an NLP algorithm to extract preoperative cannabis use status from unstructured clinical notes..* Consortium for Medical Marijuana Clinical Outcomes Research. 2023.

8. Sajdeya R, **Rouhizadeh M**, Bai C, Jugl S, Cook RL, Ison RL, Mardini MT, Winterstein AG, Price CC, Pearson TA, Tighe PJ. *Cannabis use and inhalational anesthesia maintenance: a propensity score matched retrospective cohort study.* Anesthesiology. 2023.

9. Yang KWK, Nguyen MH, Jelin A, **Rouhizadeh M**, Sobreira N, Taylor CO. *Detecting phenotypes among patients suspected of rare Mendelian disorders.* American Medical Informatics Association Informatics Summit. 2023.

10. Guan D, Li P, Fonseca V, Shi L, Ali MK, Varghese JS, Carrillo-Larco RM, **Rouhizadeh M**, Winterstein AG, Jiao T, Shao H. *Developing a machine-learning–based prediction model for diabetes duration using information from electronic health records.* Journal of Diabetes. 2023.

11. Huang W, Ahmed M, Smith S, Hasan M, **Rouhizadeh M**, Bian J, Kimmel S, Morris E, Yang L, Guo J. *Trajectories of sacubitril/valsartan adherence among Medicare beneficiaries with heart failure.* Value in Health. 2023.

12. Omaki E, Shields W, **Rouhizadeh M**, Delgado-Barroso P, Stefanos R. *Using natural language processing to understand contributing factors of pediatric falls.* Annual Meeting of American Public Health

Association. 2022.

13. Lemas DJ, **Rouhizadeh M**, Braeden L, Frank S, Wright L, Magalhães M, Xu K, Du X, Parker L, Harle C, Louis-Jaques A, Zhang B, Thompson L, Hogan WR, Modave F. *Classifying infant feeding status from clinical notes using natural language processing and machine learning.* American Medical Informatics Association Annual Symposium (AMIA). 2022.

14. Gray G, **Rouhizadeh M**, Ahumada L, Richards T, Zirikly A, Hatef E. *Application of natural language processing to identify social needs from the electronic health record's free-text notes.* American Medical Informatics Association Annual Symposium (AMIA). 2022.

15. Hatef E, **Rouhizadeh M**, Nau C, Xie F, Padilla A, Lyons LJ, Rouillard C, Abu-Nasser M, Roblin D. *Assessing the documentation of social needs in electronic health records' unstructured data: a collaboration of Johns Hopkins Health System and Kaiser Permanente.* AcademyHealth. 2021. ***Best abstract in Digital Technologies, Data and Analytics***.

16. Hatef E, **Rouhizadeh M**, Nau C, Xie F, Padilla A, Lyons LJ, Rouillard C, Abu-Nasser M, Roblin D. *A pilot study to improve the use of electronic health records for identification of patients with social determinants of health challenges: a collaboration of Johns Hopkins Health System and Kaiser Permanente.* American Medical Informatics Association Annual Symposium (AMIA). 2021.

17. Tseng E, Schwartz J, **Rouhizadeh M**, Maruthur N. *Analysis of primary care provider (PCP) EHR notes for discussions of prediabetes using natural language processing (NLP) methods.* Journal of Diabetes. 2020. No. 69, Supplement 1.

18. Taylor CO, **Rouhizadeh M**, Schiettecatte F, Levy H, Baras A, Hamosh A, Sobreira N. *Detecting phenotype descriptors in clinical notes with ClinPhen: an assessment of sensitivity in a PhenoDB patient cohort.* Annual Meeting of American Society of Human Genetics (ASHG). 2019.

19. Mulyar A, Schumacher E, **Rouhizadeh M**, Chute C, Dredze M. *Experiments with pre-trained deep neural language models for clinical NLP: concept linking and semantic similarity.* National NLP Clinical Challenges (N2C2) Workshop. 2019.

20. Dickerson L, **Rouhizadeh M**, Bowring M, Mogul D. *Use of natural language processing (NLP) to identify neurocognitive deficits in end-stage liver disease.* American Journal of Transplantation. Volume 19, Pages 29-30. 2019. ***Poster of Distinction***

21. Dickerson L, **Rouhizadeh M**, Massie A, McAdams-Demarco M, Bowring MG, Segev D, Cannon A, Guerrerio A, Chen PH, Philosophe B, Mogul D. *Language impairment in adults with end-stage liver disease: a novel application of tools from natural language processing.* Gastroenterology. Volume 156, Issue 6, Supplement 1, Page S-1369. 2019.

22. **Rouhizadeh M**, Hatef E, Dredze M, Chute C, Kharrazi H. *Identifying social determinants of health from clinical notes: a rule-based approach.* AMIA NLP working group pre-symposium. 2018.

23. Horner MS, **Rouhizadeh M**. *Homicidal and suicidal ideation as a chief complaint: considering the influence of news events using a large, de-identified electronic medical record dataset.* Society for Prevention Research. 2018.

24. **Rouhizadeh M**, Magge A, Klein A, Sarker A, Gonzalez G. *Detecting gestation period using social media data analysis.* Penn Biomedical Postdoctoral Research Symposium. 2017.

25. **Rouhizadeh M**, Schwartz HA. *Age and gender differences in self-reference in verb categories.* Mid-Atlantic Student Colloquium on Speech, Language and Learning (MASC-SLL). 2016.

26. **Rouhizadeh M**, Sproat R, van Santen J. *Computational semantic analysis of restrictive and repetitive behavior in language samples of children with autism.* International Meeting for Autism Research (IMFAR). 2015.

27. **Rouhizadeh M**, van Santen J, Sproat R, Gorman K, Heeman P, Hill AP, Bedrick S, Prud'hommeaux ET, Kiss G. *Discourse marker use in ASD and typical development.* Pacific Northwest NLP Workshop: NW-NLP. 2014.

28. **Rouhizadeh M**, van Santen J, Sproat R, Heeman P, Hill AP, Bedrick S, Prud'hommeaux E. *Children's differing patterns of discourse marker use in ASD and typical development.* International Meeting for Autism Research (IMFAR). 2014.

29. Prud'hommeaux E, **Rouhizadeh M**, Roark B, van Santen J. *Identifying unexpected and inappropriate words in ASD language samples.* International Meeting for Autism Research (IMFAR). 2013.

30. **Rouhizadeh M**, Sproat R, Coyne B. *Collecting spatial information for locations in a text-to-scene conversion system.* Pacific Northwest NLP Workshop: NW-NLP. 2012.

PRE-PRINTS

1. Zandbiglari K, Hasanzadeh HRH, Kotecha PK, Sajdeya R, Goodin AJ, Jiao T, Adiba F, Mardini MT, Bian J, **Rouhizadeh M**. A Natural Language Processing Algorithm for Classifying Suicidal Behaviors in Alzheimer's Disease and Related Dementia Patients: Development and Validation Using Electronic Health Records Data. medRxiv. 2023.
DOI: 10.1101/2023.07.21.23292976

2. Liu S, Wen A, Wang L, He H, Fu S, Miller R, Williams A, Harris D, Kavuluru R, Liu M, Abu-el-Rub N, Schutte D, Zhang R, **Rouhizadeh M**, Osborne JD, He Y, Topaloglu U, Hong SS, Saltz JH, Schaffter T, Pfaff E, Chute CG, Duong T, Haendel MA, Fuentes R, Szolovits P, Xu H, Liu H. An Open Natural Language Processing Development Framework for EHR-based Clinical Research: A case demonstration using the National COVID Cohort Collaborative (N3C). arXiv. 2021.
DOI: 10.48550/arXiv.2110.10780

3. Wang J, Pham HA, Manion F, Zhang Y, **Rouhizadeh M**. COVID-19 SignSym: a fast adaptation of a general clinical NLP tool to identify and normalize COVID-19 signs and symptoms to OMOP common data model. arXiv. 2020.
DOI: 10.48550/arXiv.2007.10286

4. Dorn R, Nobles AL, **Rouhizadeh M**, Dredze M. Examining the Feasibility of Off-the-Shelf Algorithms for Masking Directly Identifiable Information in Social Media Data. arXiv. 2020.
DOI: 10.48550/arXiv.2011.08324

5. Garibaldi BT, Fiksel J, Muschelli J, Robinson MP, **Rouhizadeh M**, Nagy P, Gray JH, Malapati H, Ghobadi-Krueger M, Niessen TM, Kim BS, Hill PM, Ahmed MS, Dobkin ED, Blanding R, Abele J, Woods B, Harkness K, Thiemann DR, Bowring MG, Shah AB, Wang MC, Bandeen-Roche K, Rosen A, Zeger SL, Gupta A. Patient trajectories and risk factors for severe outcomes among persons hospitalized for COVID-19 in the Maryland/DC region. medRxiv. 2020.
DOI: 10.1101/2020.05.24.20111864

6. Sedoc J, Wijaya D, **Rouhizadeh M**, Schwartz A, Ungar L. Deriving Verb Predicates By Clustering Verbs with Arguments. arXiv. 2017.
DOI: 10.48550/arXiv.1708.00416

EDUCATIONAL ACTIVITIES

TEACHING

University of Florida

Assistant Professor, AI in the Health Sciences Initiative 2021–present

- Introduction to Pharmaceutical Outcomes and Policy (co-instructor) Spring 2022–2024
- Measurement in Pharmaceutical Outcomes and Policy Research (co-instructor) Fall 2022, Spring 2024
- Pharmacogenomic and Genomic Data Analysis (guest lecturer) Spring 2024
- Applied Data Analysis, Interpretation, and Reporting of Findings (co-director/instructor) Fall 2023
- Introduction to Artificial Intelligence in Pharmacy (co-instructor) Fall 2023
- Practical Applications of AI to Improve Patient Safety and Quality (co-instructor) Fall 2023
- Clinical and Translational Science Institute Seminar (guest lecturer) Fall 2023
- Artificial Intelligence in Pharmacogenomics (guest lecturer) Fall 2023
- Artificial Intelligence Boot Camp (co-instructor) Fall Spring 2022-2023

Johns Hopkins University

Biomedical Informatics and Data Science (BIDS)

Adjunct Assistant Professor 2021–present
Faculty Instructor 2017–2021

- Natural Language Processing in the Health Sciences Spring 2019–2024
- Unstructured Data Mining to Address Novel Infectious Diseases Summer & Fall 2020
- NLP Methods in the Johns Hopkins Precision Medicine Analytics Platform Fall 2018, 2019
  Center of Excellence Analytics in Medicine Program (CAMP)
- Natural Language Processing in Healthcare Methodologies Spring 2020, 2021
  Leadership in Analytics and Data Science (LEADS) Program

Stony Brook University

Instructor, Department of Computer Science 2015–2017

- Symposium on Natural Language Processing for Social Science (instructor) Fall 2016
- Big Data Analytics (guest lecturer) Fall 2016
- Topics in Modern Computer Science (guest lecturer) Spring 2016
- Social Media for Computational Social Science Research (guest lecturer) Spring 2016
- Symposium on Natural Language Processing for Social Science (instructor) Fall 2015

Oregon Health & Science University

Guest Lecturer, Computer Science & Electrical Engineering 2014–2015

- Natural Language Processing Fall 2014, 2015

University of Tehran

Lecturer, English Department 2008

- General English Fall 2008

University of Applied Science and Technology

Lecturer, English Department 2007–2008

- Translation of Press Texts Fall 2008
- Contrastive Analysis Fall 2008
- Audio Video Translation Fall 2008
- English Composition Spring 2008
- Translation of Legal Texts Spring 2008
- Introduction to Linguistics Spring 2008
- Principles of Writing Winter 2008
- Translation of Economic Texts Winter 2008
- English Grammar II Winter 2008
- Reading Comprehension Fall 2007
- English Grammar I Fall 2007
- English Conversation Fall 2007

Azad University of Ramhormoz

Lecturer, Department of English 2006–2007

- General English II Spring 2007, Summer 2007
- General English I Fall 2006, Winter 2007

EDUCATIONAL PROGRAM BUILDING & LEADERSHIP

University of Florida 2021–present

College of Pharmacy AI Task Force

- Core Faculty: College of Pharmacy AI Certificate
- Course Co-Coordinator: Introduction to Artificial Intelligence in Pharmacy
- Track Lead: AI in Pharmaceutical Outcomes and Policy Research PhD/MSc Specialization

Intelligent Clinical Care Center (IC3)

- Committee Member: MS program in Artificial Intelligence in Biomedical and Health Sciences (AIBHS)

Johns Hopkins University 2017–present

Biomedical Informatics and Data Science

- Core Faculty: JHU Clinical Data Science Initiative
- Course Co-Director: Natural Language Processing in the Health Sciences
- Course Co-Director: Preparing Multidisciplinary Learners for Amazon Alexa

Stony Brook University 2017–2019

Department of Linguistics

- Advisory Board Member: Computational Linguistics Master's Program

MENTORING

University of Florida 2021–present

AI in the Health Sciences Initiative

- Hamid Hassanzadeh, PhD, Courtesy Faculty Host, *Now: Lead Data Scientist at UnitedHealth Group*
- Shobhan Kumar, PhD, Post-doctoral Advisor, *Now: Postdoc at UF*
- Muhammad Bilal, PhD, Post-doctoral Advisor, *Now: Postdoc at UF*
- Farzana Islam Adiba, MSc, PhD Advisor, *Now: PhD student at UF*
- Kimia Zandbiglari, MSc, PhD Advisor, *Now: PhD student at UF*
- Ghazal Jahanshahi, PharmD, MSc, Visiting Scholar Mentor, *Now: PhD student at Georgia Tech*
- Ruba Sajdeya, MD, Thesis Committee Member, *Now: Postdoc at UF*
- Jabed Faysal, MSc, Thesis Committee Member, *Now: PhD student at UF*
- Matthew Muschett, PharmD, Thesis Committee Member, *Now: PhD student at UF*
- Cameron Thomas, PharmD, Thesis Committee Member, *Now: PhD student at UF*
- Wenxi Huang, MSc, Thesis Committee Member, *Now: PhD student at UF*

Johns Hopkins University 2017–present

Biomedical Informatics and Data Science (BIDS)

- Nidhi Soley, MSc, Doctor of Philosophy Board Oral Exam Committee, *Now: PhD student at JHU*
- Alexandra Simpson, MD, Fellowship Committee Member, *Now: Neuroimmunology Fellow at JHU*
- Jessica Schwartz, MD, Fellowship Committee Member, *Now: Assistant Prof. of Medicine at JHU*
- Santiago Alvarez, MD, Fellowship Committee Member, *Now: Assistant Prof. of Medicine at JHU*
- Helen Ting He, MSc, Mentorship Committee Member, *Now:: PhD student at JHU*
- Kerry Smith, MSc, Mentorship Committee Member, *Now: Senior Advisor at Northwell*

Stony Brook University 2015–2017

Department of Computer Science

- Aman Raj, MSc, Thesis Committee Member, *Now: Senior Software Engineer at Google*
- Rahul Paliwal, MSc, Thesis Committee Member, *Now: Senior System Software Engineer at NVIDIA*
- Danny Bernstein, BSc, Mentorship Committee Member, *Now: Student at Harvard College*
- Rishabh Agrawal, MSc, Thesis Committee Member, *Now: Senior SDE at Amazon AWS AI*

Oregon Health & Science University 2010–2015

Computer Science & Electrical Engineering

- Margit Bowler, PhD, Mentorship Committee Member, *Now: Annotation Project Manager at Apple*
- Heather Sidener, MA, Mentorship Committee Member, *Now: Head of Clinical Medicine Unit at OHSU*

ORGANIZATIONAL ACTIVITIES

Merit Review for Funding Agencies

- NSF, Small Business Innovation Research (SBIR/STTR) - Reviewer 2023-present
- Patient-Centered Outcomes Research Institute (PCORI) - Scientific Merit Reviewer 2021-present
- UF Office of Research, Faculty Research Enhancement Program - Reviewer 2023-present
- JHU Institute for Clinical & Translational Research - Technical Reviewer 2018-present

Area Chair

- NeurIPS - Machine Learning for Health (ML4H) 2021-2023
- Association for Computational Linguistics - Cognitive Modeling & Psycholinguistics 2020

Co-Chair

- Resources and Processing of Linguistic, Para-linguistic, and Extra-linguistic Data (RaPID) 2021
- Diversity & Inclusion Subcommittee - Association for Computational Linguistics (ACL) 2020
- The 4th Pacific Northwest Regional NLP Workshop (NW-NLP) at Amazon, Seattle, WA 2016

University Governance Roles

- Intelligent Clinical Care Center (IC3) 2023-present
- UF College of Pharmacy Admissions Committee 2023-present
- AI Specialty Track Leader, UF Pharmaceutical Outcomes & Policy 2022-present
- UF College of Pharmacy AI Task Force 2021-present
- UF Research Computing Advisory Committee Member 2011-present

Conference Organizing Committees

- The 11th Conference on Information and Knowledge Technology (IKT2020) 2020
- Diversity & Inclusion - North American Association for Computational Linguistics 2019
- Conference of the International Speech Communication Association (Interspeech) 2012
- The 49th Annual Meeting of the Association for Computational Linguistics (ACL-HLT) 2011

Special Interest Group Leadership

- Capturing Socio-Behavioral Determinants of Health Data in Johns Hopkins EHR 2021
- IEEE Spoken Language Technology Workshop (SLT) 2014

Editorial Board Appointments

- Nature – Humanities and Social Sciences Communications 2022-present
- Frontiers in Drug Safety and Regulation 2021-present

Journal Peer Review Activities

- Journal of Healthcare Informatics Research 2022-present
- Nature Humanities and Social Sciences Communications 2022-present
- Nature Partner Journals - Digital Medicine 2022-present
- Journal of Medical Internet Research 2021-present
- Journal of Biomedical Informatics 2021-present
- Journal of Natural Language Engineering 2020-present
- Journal of the American Medical Informatics Association 2020-present
- Patterns - Cell Press 2020-present
- PLOS ONE 2019-2021
- Information Processing and Management (Elsevier IPM) 2018-present
- ACM Transactions on Asian Language Information Processing (TALLIP) 2013-2020
- Language Resources and Evaluation Journal (LREV) 2012-2016

Scientific Program Committees

- Resources and Processing of Linguistic, Para-linguistic, and Extra-linguistic Data 2022-2024
- Language Resources and Evaluation (LREC) 2018, 2020, 2024
- American Medical Informatics Association (AMIA) Annual Symposium 2019-2023
- AI Techniques in Interaction-centric Autism Research and Diagnosis 2023
- AcademyHealth Annual Research Meeting (ARM) 2021-2022
- AMIA NLP Working Group 2022-2023
- Association for Computational Linguistics (ACL-IJCNLP) 2020-2021
- North American Association for Computational Linguistics (NAACL) 2016, 2018-2019, 2021
- Annual Meeting of the Cognitive Science Society (CogSci) 2020-2021
- Computational Linguistics and Clinical Psychology (CLPsych) 2017-2021
- Empirical Methods in Natural Language Processing (EMNLP) 2017-2020
- IEEE Signal Processing Society - Acoustics, Speech, and Signal Processing (ICASSP) 2020
- American Medical Informatics Association (AMIA) Informatics Summit 2018-2020
- Negative Results in NLP 2020
- Computational Natural Language Learning (CoNLL) 2019
- Speech and Language Processing for Assistive Technologies (SLPAT) 2019
- Abusive Language Online (ALW) 2017-2019
- Pacific Northwest Regional NLP Workshop (NW-NLP) 2014, 2016, 2018
- IEEE Human-Centered Computational Sensing (HCCS) 2018
- Pacific Symposium on Biocomputing (PSB) 2018
- Social Media Mining for Health Applications (SMM4H) 2017
- Biomedical Natural Language Processing (BioNLP) 2017
- International Conference on Computational Linguistics (COLING) 2016

Professional Society Memberships

- American Association of College of Pharmacy (AACP) 2021-present
- Open Health Natural Language Processing Consortium (OHNLP) 2020-present
- NIH National COVID Cohort Collaborative (N3C) 2020-present
- Health Level Seven International (HL7) Working Group 2019-present
- Observational Health Data Sciences and Informatics (OHDSI) 2017-present
- American Medical Informatics Association (AMIA) NLP Working Group 2017-present
- The New York Academy of Sciences 2015-present
- Association for Computing Machinery (ACM) 2013-present
- Association for Computational Linguistics (ACL) 2010-present
- Global WordNet Association (GWA) 2007-present
- International Society for Autism Research (INSAR) 2013-2015
- International Speech Communication Association (ISCA) 2014-2015

NATIONAL COLLABORATIVES AND SYSTEM INNOVATION

Observational Health Data Sciences and Informatics (OHDSI)

- Co-leading a multi-institutional OHDSI collaboration to evaluate NLP pipelines for robust identification of social determinants of health from clinical notes.
- Co-leading efforts to enhance clinical concept extraction and normalization to the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM).
- Collaborating with the OHDSI community on the next generation of OHDSI OMOP CDM for standardizing NLP-derived clinical data, contributing to the development of the new OMOP Note_NLP table.

Open Health Natural Language Processing Consortium (OHNLP)

- Advanced open-source development for scientific progress, contributing Python tools for identifying signs and symptoms, social determinants of health, and rare diseases in vulnerable populations.

NIH National COVID Cohort Collaborative (N3C)

- Directed NLP research in the COVID-19 Precision Medicine Analytics Platform Registry (JH-CROWN), collaborating with the National COVID Cohort Collaborative, Johns Hopkins, Mayo Clinic, and UTHealth to develop tools for extracting COVID-19 symptoms from clinical notes and normalizing them into the OMOP CDM.
- Expanded tool development to address neurological manifestations of COVID-19 and contributed to a project for identifying symptom severity, onset, and duration of Post-Acute Sequelae of SARS-CoV-2 (PASC).

American Medical Informatics Association (AMIA) NLP Working Group

- Actively served in roles as instructor, event organizer, doctoral consortium faculty, and panelist, raising the profile of my work and institution.

RECOGNITION

INVITED TALKS

- International Society for Pharmacoepidemiology (ISPE) Meeting, *Natural Language Processing for the Analysis of Unstructured Clinical Data*, Orlando, FL, April 2024.
- American Pharmacists Association (APhA) Annual Meeting, *Elevating Pharmacy Practice with Artificial Intelligence (AI)*, Orlando, FL, March 2024.
- DigitalHealth.Rx Summit, *Embracing Digital Health Innovations to Advance Pharmacy Practice*, Orlando, FL, March 2024.
- University of Florida, College of Pharmacy Artificial Intelligence PharmTalks, *Leveraging Advanced Natural Language Processing Methods in Pharmacy Research and Practice: An Overview from a Real-World Data Perspective*, Gainesville, FL, April 2023.
- Rita Kobb Nursing and Health Informatics Symposium, *Investigating AI and Natural Language Processing Methods for Extracting Social Determinants of Health from Unstructured Data in Electronic Medical Records*, Gainesville, FL, Feb 2023.
- Johns Hopkins All Children's Hospital, *Exploring Deep Learning and AI Techniques for Extracting Information from Unstructured Data in Electronic Medical Records*, St. Petersburg, FL, Feb 2023.
- University of Florida, AI2HEAL DATATHON, *AI and NLP Methods to Extract Socio-Behavioral Data from Unstructured Electronic Medical Records*, Gainesville, FL, Jan 2023.
- American Medical Informatics Association, *Identifying Social Determinants of Health (SDoH) Using Natural Language Processing*, Washington, DC, Nov 2022.
- Loyola University Chicago Health Informatics Seminar Series, *Exploring Deep Learning and NLP Techniques in Electronic Health Record Data Mining*, Chicago, IL, Dec 2021.
- The University of Tennessee Health Science Center, *AI and Natural Language Processing Methods for Extracting Clinical Concepts and Identifying Social Determinants of Health from Clinical Notes*, Memphis, TN, Dec 2021.
- N3C NLP Working Group, *Leveraging Big Data for Population Health Insights: Progressing NLP Methods to Identify Social-Behavioral Risk Factors in EHR Text*, Nov 2021.

- OHDSI NLP Working Group, *Utilizing Big Data in Population Health: Enhancing NLP Approaches for Extracting Social-Behavioral Risk Factors from EHR Free Text*, Oct 2021.
- University of Florida, Department of Health Outcomes and Biomedical Informatics, *Deep Learning and Natural Language Processing Methods for Mining Electronic Medical Records*, Gainesville, FL, June 2021.
- Mayo Clinic, Department of AI and Informatics, *Deep Learning and Natural Language Processing Methods for Mining Electronic Medical Records*, Rochester, MN, June 2021.
- The University of Texas Health Science Center at Houston, *Deep Learning and NLP Methods for Mining Electronic Medical Records*, Houston, TX, May 2021.
- Columbia University Irving Medical Center, *Natural Language Processing in the Health Sciences*, New York, NY, April 2021.
- Columbia University Irving Medical Center, *Deep Learning and NLP Methods for Mining Electronic Medical Records*, New York, NY, April 2021.
- Johns Hopkins University, Symposium on Capturing Socio-Behavioral Determinants of Health Data, *Natural Language Processing Methods to Extract Socio-Behavioral Data from Unstructured EHR*, Baltimore, MD, March 2021.
- Johns Hopkins Medicine, Department of General Internal Medicine, *Using Natural Language Processing Methods to Identify Discussions of Prediabetes in EHR Notes*, Baltimore, MD, December 2020.
- Johns Hopkins University, Center for Language and Speech Processing, *Deep Learning and NLP Methods for Mining Electronic Medical Records*, Baltimore, MD, October 2020.
- Chalmers University of Technology, Department of Computer Science and Engineering, *How Will AI Influence the Engineering Profession?*, Gothenburg, Sweden, December 2019.
- Chalmers University of Technology, Department of Computer Science and Engineering, *Deep Learning and NLP Methods in Healthcare*, Gothenburg, Sweden, December 2019.
- American Medical Informatics Association, Public Health Informatics Working Group, *NLP and Deep Learning Models for Extracting Housing Challenges from EHR*, Washington, DC, November 2019.
- NYU Langone Health, *Deep Learning and NLP Methods for Mining Electronic Medical Records*, New York, NY, October 2019.
- Johns Hopkins University, Biomedical Informatics and Data Science Grand Rounds, *Applications of Natural Language Processing in Identifying Language Disorders*, Baltimore, MD, December 2018.
- National Research Council of Canada, *Distributional Semantic Methods for the Characterization of Semantic and Pragmatic Facets of Language*, Ottawa, ON (remote presentation), February 2018.
- Johns Hopkins University, Center for Language and Speech Processing, *Distributional Semantic Methods for Characterization of Atypical Language In Autism*, Baltimore, MD, July 2016.
- University of Pennsylvania, Positive Psychology Center, *Computational Methods for Categorization of Disordered Language*, Philadelphia, PA, April 2016.
- Stony Brook University, Department of Linguistics, *Computational Semantic Analysis of Atypical Language in Autism*, Stony Brook, NY, March 2016.

INVITED PANEL DISCUSSIONS

Invited Panel Discussions

- Rita Kobb Nursing and Health Informatics Symposium, *Investigating AI and Natural Language Processing Methods for Extracting Social Determinants of Health from Unstructured Data in Electronic Medical Records*, Gainesville, FL, February 2023.

- American Medical Informatics Association NLP Working Group, *Leveraging NLP Methods for Exploring Social and Behavioral Determinants of Health*, Washington, DC, November 2022.
- American Medical Informatics Association NLP Working Group, *Faculty Panel of AMIA NLP-WG Doctoral Consortium*, San Diego, CA, November 2021.
- Johns Hopkins University, Symposium on Capturing Socio-Behavioral Determinants of Health Data, *Natural Language Processing Methods to Extract Socio-Behavioral Data from Unstructured EHR*, Baltimore, MD, March 2021.
- American Medical Informatics Association, Public Health Informatics Working Group, *Artificial Intelligence Methods for Extracting Socio-economic Challenges from EHR*, Washington, DC, November 2019.
- IEEE Spoken Language Technology Workshop, *Innovative Assistive Technologies for Speech and Language Disorders*, South Lake Tahoe, NV, December 2014.

ORAL & POSTER PRESENTATIONS

- American Medical Informatics Association Annual Symposium (AMIA), New Orleans, Louisiana, USA, November 2023, Poster Presentation: *Preoperative Cannabis Use Status Ascertainment Using Natural Language Processing Methods and Coded Data.* ***Best Poster Award Nominee***
- American Medical Informatics Association Annual Symposium (AMIA), New Orleans, Louisiana, USA, November 2023, Poster Presentation: *Deep Learning and Natural Language Processing Methods for Characterizing Pediatric Falls Events.*
- American Medical Informatics Association Annual Symposium (AMIA), New Orleans, Louisiana, USA, November 2023, Poster Presentation: *Development and Validation of a Natural Language Processing Algorithm for Extracting Preoperative Cannabis Use Status Information from Unstructured Narrative Clinical Notes.*
- Consortium for Medical Marijuana Clinical Outcomes Research, Orlando, FL, June 2023, Poster Presentation: *Developing and validating an NLP algorithm to extract preoperative cannabis use status from unstructured clinical notes..*
- Anesthesiology Conference, San Francisco, CA, April 2023, Poster Presentation: *Cannabis Use and Inhalational Anesthesia Maintenance: A Propensity Score Matched Retrospective Cohort Study.*
- American Medical Informatics Association Annual Symposium (AMIA), New Orleans, Louisiana, USA, May 2023, Poster Presentation: *Developing a Machine-Learning–Based Prediction Model for Diabetes Duration Using Information from Electronic Health Records.*
- American Medical Informatics Association Informatics Summit, Seattle, WA, USA, March 2023, Poster Presentation: *Detecting Phenotypes Among Patients Suspected of Rare Mendelian Disorders.*
- Professional Society for Health Economics and Outcomes Research (ISPOR), Boston, MA, March 2023, Oral Presentation: *Trajectories of Sacubitril/Valsartan Adherence Among Medicare Beneficiaries with Heart Failure.*
- American Medical Informatics Association Informatics Summit, Seattle, WA, February 2023, Poster Presentation: *Detecting Phenotypes Among Patients Suspected of Rare Mendelian Disorders.*
- American Public Health Association, Boston, MA, November 2022, Poster Presentation: *Using Natural Language Processing to Understand Contributing Factors of Pediatric Falls.*
- American Medical Informatics Association, Washington, DC, November 2022, Oral Presentation: *Application of Natural Language Processing to Identify Social Needs from The Electronic Health Record's Free-Text Notes.*

- American Medical Informatics Association, Washington, DC, November 2022, Poster Presentation: *Classifying Infant Feeding Status from Clinical Notes Using Natural Language Processing and Machine Learning.*
- Anesthesiology Conference, San Francisco, CA, Poster Presentation: *Cannabis Use and Inhalational Anesthesia Maintenance: A Propensity Score Matched Retrospective Cohort Study.*
- AMIA-affiliated Workshop on Challenges in NLP for Clinical Data, Washington, DC, November 2019, Oral Presentation: *Experiments with Pre-Trained Deep Neural Language Models for Clinical NLP: Concept Linking and Semantic Similarity.*
- Department of Epidemiology, Biostatistics, and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, November 2017, Poster Presentation: *Mining Pregnancy-Related Health Information from Social Media.*
- The 16th Annual Biomedical Postdoctoral Research Symposium, University of Pennsylvania, Philadelphia, PA, October 2017, Poster Presentation: *Detecting Gestation Period Using Social Media Data Analysis.*
- Conference on Empirical Methods in Natural Language Processing (EMNLP), Austin, TX, November 2016, Poster Presentation: *Using Syntactic and Semantic Context to Explore Psychodemographic Differences in Self-reference.*
- The 4th Pacific Northwest NLP Workshop (NW-NLP), Amazon, Seattle, WA, May 2016, Oral Presentation: *Measuring Idiosyncratic Interests in Children with Autism Spectrum Disorder.*
- 4th Pacific Northwest NLP Workshop (NW-NLP), Amazon, Seattle, WA, May 2016, Poster Presentation: *Quantifying Restrictive and Repetitive Interest in Conversations of Autistic Children.*
- Mid-Atlantic Student Colloquium on Speech, Language and Learning (MASC-SLL), Philadelphia, PA, April 2016, Poster Presentation: *Age and Gender Differences in Self-reference in Verb Categories.*
- Computational Linguistics and Clinical Psychology Workshop (CLPsych) at NAACL, Denver, CO, June 2015, Oral Presentation: *Similarity Measures for Quantifying Restrictive and Repetitive Behavior in Conversations of Autistic Children.*
- International Meeting for Autism Research, Salt Lake City, UT, May 2015, Poster Presentation: *Computational Semantic Analysis of Restrictive and Repetitive Behavior in Language Samples of Children with Autism.*
- IEEE Spoken Language Technology Workshop (SLT), South Lake Tahoe, NV, December 2014, Poster Presentation: *Computational Analysis of Trajectories of Linguistic Development in Autism.*
- Computational Linguistics & Clinical Psychology Workshop (CLPsych) at ACL, Baltimore, MD, June 2014, Oral Presentation: *Detecting Linguistic Restricted Interests in Autism Using Distributional Semantic Models.*
- Center for Spoken Language Understanding, Oregon Health & Science University, Portland, OR, June 2014, Poster Presentation: *Detecting Linguistic Restricted Interests in Autism.*
- International Meeting for Autism Research (IMFAR), Atlanta, GA, May 2015, Poster Presentation: *Children's Differing Patterns of Discourse Marker Use in ASD and Typical Development.*
- Center for Spoken Language Understanding, Oregon Health & Science University, Portland, OR, May 2013, Oral Presentation: *Distributional Semantic Applications in the Evaluation of Disordered Speech.*
- Annual Conference of the Institute on Development & Disability, Oregon Health & Science University, Portland, OR, May 2013, Poster Presentation: *Distributional Semantic Models for the Evaluation of Disordered Language.*
- Annual Conference of the Institute on Development & Disability, Oregon Health & Science University, Portland, OR, May 2013, Poster Presentation: *Identifying Unexpected and Inappropriate Words in ASD Language Samples.*

- Center for Spoken Language Understanding, Oregon Health & Science University, Portland, OR, February 2012, Oral Presentation: *Distributional Semantic Applications in the Evaluation of Disordered Speech.*
- Center for Spoken Language Understanding, Oregon Health & Science University, Portland, OR, October 2012, Oral Presentation: *Knowledge Representation and Annotation for a Text-To-Scene System.*
- Language and Speech Processing Group, Department of Computer Science, Columbia University, New York, NY, August 2012, Oral Presentation: *Crowd-sourcing Semantic Information Collection.*
- 2nd Pacific Northwest NLP Workshop (NW-NLP), Microsoft Research, Redmond, WA, May 2012, Poster Presentation: *Collecting Spatial Information for Locations in a Text-to-Scene Conversion System.*
- Center for Spoken Language Understanding, Oregon Health & Science University, Portland, OR, December 2010, Oral Presentation: *Collecting Semantic Data by Mechanical Turk for the Lexical Knowledge Resource of a Text-to-Picture Generating System.*
- 2nd Workshop on Natural Language Processing, Shahid Beheshti University, Tehran, Iran, May 2007, Oral Presentation: *Introduction to Lexical Resources and Ontologies.*
- 7th Conference on Iranian Linguistics, Allameh Tabatabaei University, Tehran, Iran, December 2007, Oral Presentation: *Designing the WordNet for Persian Verbs.*

LANGUAGES

• English	Highly proficient
• Persian	Native
• Arabic	Intermediate
• Old Persian	Basic
• Middle Persian	Basic
• Avestan	Basic
• Italian	Elementary

COMPUTER SKILLS

Machine Learning & Deep Learning

• PyTorch	• TensorFlow	• scikit-learn
• Keras	• OpenCV	• CUDA
• Pandas	• Matplotlib	• Seaborn
• NumPy	• SciPy	

Natural Languages Processing

• spaCy	• cTAKES	• MetaMap
• scispaCy	• flairNLP	• SyntaxNet
• NLTK	• FastText	• Stanford CoreNLP
• Gensim	• Word2Vec	• ARK TweetNLP
• Lucene	• OpenFST	• Lextools
• Praat	• Thrax	• HTK
• Moses	• GIZA++	• E-Prime

Languages & Databases

- Python
- R
- C
- C++
- Prolog
- Java
- MySQL
- PostgreSQL
- UNIX Shell Scripting
- HBase
- MongoDB
- SQLite

Frameworks & Technologies

- AWS
- Azure
- Slurm
- Spark
- Hadoop
- MapReduce
- PHP
- XML
- CSS
- HTML
- Jira
- Condor