Anthony Rios | Assistant Professor

North Paseo Building — One UTSA Circle — San Antonio, TX 78249
☎ (210) 458-6303 • ✉ anthony.rios@utsa.edu • www.anthonyrios.net

Research Interests

Machine Learning; Natural Language Processing; Neural Networks; Biomedical Informatics; Computational Social Science

Education

University of Kentucky
Ph.D. in Computer Science
Advisor: Ramakanth Kavuluru, Ph.D.
Co-Advisor: Miroslaw Truszczynski, Ph.D. *2012–2018*

Georgetown College
B.S. in Computer Science *2007–2011*

Professional Experience

University of Texas at San Antonio
Assistant Professor in Information Systems and Cyber Security *2018–present*

Associate Director of the Security and Analytics *2021–present*

National Institute of Health (NCBI/NLM/NIH)
Summer Research Fellow
Supervisor: Zhiyong Lu, Ph.D. *Summer 2017*

University of Kentucky
Graduate Research Assistant *2013–2018*

Awards

2022 – NSF CAREER Award

2021 – JAMIA Editor's Choice *and* Featured Research Paper

2019 – UTSA's AI Summit Special Recognition Award

2017 – Best poster, Annual Commonwealth Computational Summit

2017 – Ranked 1st in the BioCreative text mining chemical-protein interactions (CHEMPROT) shared task

2017 – Ranked 2nd in the shared task on classification of medication intake messages on Twitter for online pharmacovigilance (at Social media mining for health workshop at AMIA)

2017 – NIH Intramural Research Training Award (IRTA)

2016 – Ranked 3rd in the CEGS NGRID shared task on predicting psychiatric symptom severity scores

based on clinical notes (RDoC for Psychiatry workshop at AMIA)

2016 – Graduate School Travel Grant, University of Kentucky

2015 – Thaddeus B. Curtz Memorial Scholarship, University of Kentucky

2015 – Best paper nomination, IEEE International conference on healthcare informatics, IEEE ICHI 2015.

2015 – Ranked 2nd in the Annual BioASQ Semantic Indexing Challenge, Task A (Batch 2)

2014 – Distinguished poster nomination, American Medical Informatics Assoc. (AMIA) Annual Symposium

2011 – Outstanding Senior in Computer Science, Georgetown College

Research Funding

I have currently received a total of $661,825 in research funding.

CAREER: Learning and Using Community-Driven Natural Language Processing Models.
National Science Foundation. CISE: IIS. 06/01/2022-05/30/2027. $551,575
PI: *Anthony Rios* (Sole)
Status: **Awarded**

CRII: SCH: A Computational Framework for Fair Public Health-Related Decisions.
National Science Foundation. CISE: IIS. 04/01/2020-03/31/2023. $174,797
PI: *Anthony Rios* (Sole)
Status: **Awarded**

Machine Learning-centric Cyber Threat Intelligence and Hunting for IoT Systems.
National Security Agency (NSA/DoD). 08/27/2021-08/26/2024. $487,028
PI: *Anthony Rios*, CoPI(s): Raymond Choo and Glenn Dietrich
Status: **Awarded**

Peer-Reviewed Journal Publications

Impact factors are from the year the article was published. For current impact factors see my research statement. See current citation information via Google Scholar at `https://scholar.google.com/citations?user=KJr3ptUAAAAJ`.

Papers published by students under my supervision are marked with an asterisk (*).

[J 1] **The Risk of Racial Bias while Tracking Influenza-Related Content on Social Media using Machine Learning.**
Brandon Lwowski* and *Anthony Rios*.
Journal of the American Medical Informatics Association (JAMIA), 2021.
(Impact Factor: **4.112**, **Editor's Choice**, **Featured Article**)

[J 2] **Cross-registry neural domain adaptation to extract mutational test results from pathology reports.**
Anthony Rios, Eric Durbin, Isaac Hands, Susanne Arnold, Darshil Shah, Stephen Schwartz, Bernardo Goulart, and Ramakanth Kavuluru.

Journal of Biomedical Informatics, 2019.
(Impact Factor: **2.950**)

[J 3] **Neural Transfer Learning for Assigning Diagnosis Codes to EMRs.**
Anthony Rios and Ramakanth Kavuluru.
Artificial Intelligence in Medicine, 2019.
(Impact Factor: **3.574**)

[J 4] **Data and systems for medication-related text classification and concept normalization from Twitter: Insights from the Social Media Mining for Health (SMM4H) 2017 shared task.**
Abeed Sarker, Maksim Belousov, Jasper Friedrichs, Kai Hakala, Svetlana Kiritchenko, Farrokh Mehryary, Sifei Han, Tung Tran, *Anthony Rios*, Ramakanth Kavuluru, Berry de Bruijn, Filip Ginter, Debanjan Mahata, Saif M Mohammad, Goran Nenadic, and Graciela Gonzalez-Hernandez.
Journal of the American Medical Informatics Association (JAMIA), 2018.
(Impact Factor: **4.270**)

[J 5] **Extracting chemical-protein relations with ensembles of SVM and deep learning models.**
Yifan Peng, *Anthony Rios*, Ramakanth Kavuluru, and Zhiyong Lu.
Database: The Journal of Biological Databases and Curation, 2018.
(Impact Factor: **3.290**)

[J 6] **Generalizing Biomedical Relation Classification with Neural Adversarial Domain Adaptation.**
Anthony Rios, Ramakanth Kavuluru, and Zhiyong Lu.
Bioinformatics, 2018.
(Impact Factor: **7.307**)

[J 7] **Ordinal Convolutional Neural Networks for Predicting RDoC Positive Valence Psychiatric Symptom Severity Scores.**
Anthony Rios and Ramakanth Kavuluru.
Journal of Biomedical Informatics, 2017.
(Impact Factor: **2.753**)

[J 8] **An Empirical Evaluation of Supervised Learning Approaches in Assigning Diagnosis Codes to Electronic Medical Records.**
Ramakanth Kavuluru, *Anthony Rios*, and Yuan Lu.
Artificial Intelligence in Medicine, 2015.
(Impact Factor: **2.009**)

Peer-Reviewed Conference Publications

For computer science publication venues, conference proceedings are highly selective, and are considered prestigious—sometimes **more prestigious than journals**. Therefore, highly-selective conferences are generally considered the most respected publication venues within certain communities such as natural language processing (e.g., ACL, NAACL, EMNLP, and AAAI). See csrankings.com and Google scholar for lists of top venues. See current citation information via Google Scholar at `https://scholar.google.com/citations?user=KJr3ptUAAAAJ`.

Papers published by students under my supervision are marked with an asterisk (*).

[C 1] **Turning Stocks into Memes: A Dataset for Understanding How Social Communities Can Drive Wall Street.**
Richard Alvarez*, Paras Bhatt*, Xingmeng Zhao*, and *Anthony Rios*.
To Appear In: Proceedings of the International AAAI Conference on Web and Social Media, ICWSM, 2022.

[C 2] **Assigning ICD-O-3 Codes to Pathology Reports using Neural Multi-Task Training with Hierarchical Regularization.**
Anthony Rios, Eric Durbin, Isaac Hands and Ramakanth Kavuluru.
In: Proceedings of the ACM Conference on Bioinformatics, Computational Biology, and Health Informatics, BCB 2021.

[C 3] **Detecting Machine-Generated Text by Characterizing Linguistic Accommodation in Human-Machine Interactions.**
In: Findings of the Association of Computation Linguistics, ACL: Findings 2021.
Paras Bhatt* and *Anthony Rios*.
(Acceptance Rate: **36.2%**)

[C 4] **An Empirical Study of the Downstream Reliability of Pre-Trained Word Embeddings.**
Anthony Rios and Brandon Lwowski.
In: Proceedings of the 28th International Conference on Computational Linguistics, COLING 2020.
(Acceptance Rate: **32.9%**)

[C 5] **FuzzE : Fuzzy Fairness Evaluation of Offensive Language Classifiers on African-American English.**
Anthony Rios.
In: Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020.
(Acceptance Rate: **20%**)

[C 6] **Complex contagions of information diffusion across social networking platforms.**
Rachael Ruizhu Xiong, Charles Zhechao Liu, Kim-Kwang Raymond Choo, and *Anthony Rios*.
In: Proceedings of the Americas Conference on Information Systems, AMCIS 2019.

[C 7] **Few-Shot and Zero-Shot Multi-Label Learning for Structured Label Spaces.**
Anthony Rios and Ramakanth Kavuluru.
In: Proceedings of the Conference on Empirical Methods in Natural Language Processing, EMNLP 2018.
(Acceptance Rate: **25%**)

[C 8] **EMR Coding with Semi-Parametric Multi-Head Matching Networks.**
Anthony Rios and Ramakanth Kavuluru.
In: Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1, Long Papers, NAACL 2018.
(Acceptance Rate: **32%**)

[C 9] **Automatic Assignment of Non-Leaf Medical Subject Headings to Biomedical Articles.**
Ramakanth Kavuluru and *Anthony Rios*.
In: Proceedings of the American Medical Informatics Association Symposium, AMIA 2015.

[C 10] **Analyzing the Moving Parts of a Large-Scale Multi-Label Text Classication Pipeline: Experiences in Indexing Biomedical Articles.**
Anthony Rios and Ramakanth Kavuluru.
In: Proceedings of the IEEE International Conference on Healthcare Informatics, ICHI 2015.
(Acceptance Rate: **28%**, **Best Paper Finalist**)

[C 11] **Convolutional Neural Networks for Biomedical Text Classification: Application in Indexing Biomedical Articles.**
Anthony Rios and Ramakanth Kavuluru.
In: Proceedings of the ACM Conference on Bioinformatics, Computational Biology, and Health Informatics, BCB 2015.
(Acceptance Rate: **34%**)

[C 12] **A Knowledge-Based Collaborative Clinical Case Mining Framework.**
Ramakanth Kavuluru, *Anthony Rios*, Brandom Kulengowski, and Patrick McNamara.
In: Proceedings of the American Medical Informatics Association Symposium, AMIA 2014.

[C 13] **Supervised Extraction of Diagnosis Codes from EMRs: Role of Feature Selection, Data Selection, and Probabilistic Thresholding.**
Anthony Rios and Ramakanth Kavuluru.
In: Proceedings of the IEEE International Conference on Healthcare Informatics, ICHI 2013.
(Acceptance Rate: **30%**)

[C 14] **A Multi-Label Classication Approach to Coding Cancer Information Service Chat Transcripts.**
Anthony Rios, Robin Vanderpool, P. Shaw, and Ramakanth Kavuluru.
In: Proceedings of the Twenty-Sixth International Conference of the Florida Artificial Intelligence Research Society, FLAIRS 2013.

Peer-Reviewed Workshop Publications

Papers published by students under my supervision are marked with an asterisk (*).

[W 1] **A Self-supervised Approach for Semantic Indexing in the Context of the COVID-19 Pandemic**.
Nima Ebadi, *Anthony Rios*, and Paul Rad
Under Review In: BioNLP, ACL 2022.

[W 2] **UTSA NLP at SemEval-2022 Task 4: An Exploration of Simple Ensembles of Transformers, Convolutional, and Recurrent Neural Networks**.
Xingmeng Zhao* and *Anthony Rios*
Under Review in: SemEval, NAACL 2022.

[W 3] **Quantifying 60 Years of Gender Bias in Biomedical Research with Word Embeddings**.
Anthony Rios, Reenam Joshi*, and Hejin Shin*
In: Proceedings of BioNLP, ACL 2020.
(Acceptance Rate: **30%**)

[W 4] **How Many Users Are Enough? Exploring Semi-Supervision and Stylometric Features to Uncover a Russian Troll Farm**.
Nayeema Nasrin*, Raymond Choo, Myung Ko, and *Anthony Rios*
In: Proceedings of the Workshop on NLP for Information Freedom, EMNLP 2019.

[W 5] **Predicting Psychological Health from Childhood Essays with Convolutional Neural Networks for the CLPsych 2018 Shared Task**.
Anthony Rios, Tung Tran, and Ramakanth Kavuluru.
In: Proceedings of the Workshop on Computational Linguistics and Clinical Psychology, NAACL 2018.

[W 6] **Team UKNLP: Detecting ADRs, Classifying Medication In-take Messages, and Normalizing ADR Mentions on Twitter.**
Sifei Han, Tung Tran, *Anthony Rios*, Ramakanth Kavuluru.
In: Proceedings of the Social Media Mining for Health Applications Workshop, AMIA 2017.

[W 7] **Chemical-protein relation extraction with SVM, CNN, RNN and ensemble systems.**
Yifan Peng, *Anthony Rios*, Ramakanth Kavuluru, and Zhiyong Lu.
In: Proceedings of the BioCreative Workshop, 2017.

[W 8] **Extracting Drug-Drug Interactions with Word and Character-Level Recurrent Neural Networks.**
Ramakanth Kavuluru, *Anthony Rios*, and Tung Tran.
In: Proceedings of the Workshop on Healthcare Knowledge Discovery and Management, ICHI 2017.

Theses

[T 1] **Deep Neural Networks for Multi-Label Text Classification: Application to Coding Electronic Medical Records.**
Anthony Rios.
University of Kentucky, 2018.

Invited Talks and Other Workshop Presentations

Invited talk titled "Bias in Biomedical NLP" at Stanford Emergency Medicine Innovation Symposium.

"The good the bad and the ugly of Biomedical NLP" guest lecture for the Biomedical NLP class at Weill Cornell Medicine.

Generalizing Biomedical Relation Extraction with Neural Adversarial Domain Adaptation. Bluegrass Data Science Group, 2018.

Adversarial Discriminative Domain Adaptation for Extracting Protein-Protein Interactions from Text. Annual Commonwealth Computational Summit, 2017. (**Best Poster Award**)

Ordinal Convolutional Neural Networks for RDoc Classification. CEGS N-GRID Shared-Tasks and Workshop on Challenges in Natural Language Processing for Clinical Data, 2016.

Convolutional Neural Networks for Biomedical Text Classification: Applications in Indexing Biomedical Articles, Keeping Current, University of Kentucky, and Lexmark International 2016.

Multi-label Collective Classification, Keeping Current, University of Kentucky, Department of Computer Science 2014.

Data Science Workflow with IPython Notebook, Keeping Current, University of Kentucky, Department of Computer Science 2014.

Sick Jump: Maximizing Vertical Air to Optimize Tricks on a Half-pipe. Kentucky Section of the MAA Annual Meeting, Eastern Kentucky University, 2011.

Industrial Experience

MAKETIME Inc. (Now Xometry)
Data Scientist *2016–2018*

Lexmark International
Software Engineer Intern/Co-op *2010–2013*

Corevalus Systems LLC.
Software Engineer Intern *Spring 2010*

Professional Memberships and Activities

Reviewer for the Journal of Biomedical Informatics (JBI), the Information Sciences Journal, Bioinformatics, Plos ONE, American Medical Informatics Association Annual Symposium (AMIA), The Journal of the American Medical Informatics Association Annual Symposium (JAMIA), IEEE TrustCom, IEEE Access, IEEE Journal of Biomedical and Health Informatics, ACL, EMNLP, NAACL CoNLL

Member of the American Medical Informatics Association (AMIA), Association for Computational Linguistics (ACL), and the American Heart Association, Association for Computer Machinery (ACM), Association for the Advancement of Artificial Intelligence (AAAI)

Session chair for AAAI 2020 and COLING 2020.

Scientific Program Committee (SPC) for the 2021 AMIA Informatics Summit.

Local Arrangements Chair for IEEE ISI 2021.

Teaching Experience

University of Texas at San Antonio **Instructor**
Data Foundations *Fall 2018, Spring 2019*

University of Texas at San Antonio **Instructor**
Introduction to Natural Language Processing *Fall 2019*

University of Texas at San Antonio **Guest Lecturer**
Introduction to Artificial Intelligence (Computer Science) *Fall 2018*

University of Kentucky **Guest Lecturer**
Biomedical Natural Language Processing *Fall 2016*

Georgetown College **Tutor**
Computer Science Peer Tutor *2010–2011*

Ph.D. Research Mentor

Rachael Xiong, (Spring 2019-present), IS Ph.D.

Nayeema Nasrin, (Fall 2018-present), IS Ph.D.

Xingmeng Zhao, IS Ph.D.

Yuhe Ding, (Fall 2018), IS Ph.D.

Masters Committee Member

Lee Boyd (Fall 2019) CS, Chair: Weining Zhang

Manan Bhagat (Spring 2019) ECE, Chair: Miltos Alamaniotis

Independent Study

Nayeema Nasrin, (Fall 2019), Programming for data science

Hejin Shin, (Summer 2019), Temporal text analysis

Service to the University of Texas at San Antonio

Information Systems PhD Committee, 2018–present

Information Systems Faculty Search Committee, 2018–present

Information Systems Undergraduate Programs Committee, 2018–present

Selected Open Source Software `https://github.com/AnthonyMRios`

pyMetaMap – Python interface for the widely used named entity recognition tool (MetaMap) by the National Library of Medicine.

pyClausIE – Python interface for the OpenIE tool ClausIE.

relation-extraction-rnn – Python package including a Bi-directional LSTM method for relation extraction.

LEML – Python package for a matrix factorization-based method for extreme mulit-label classification.

advRelCNN – Python package including a method for unsupervised domain adaptation for relation classification.