Research Collaboration Seed Funding

Proposal Cover Page

Please complete the information in the shaded areas below.
Submit this cover page with your proposal and supporting documents as one document in MS Word or PDF format to: swilmoth@uark.edu

Proposal Title: A prototype to understand spatial spillovers of clinical outcomes for better policy construction.

Submitting Principal Investigator Information:

Name: Aranyak Goswami
Phone: 479-575-4351
Email: garanyak@uark.edu
Department: Department of Agriculture (AFLS)

Collaborator Information:

Role	Name	Department
Co-PI	Abhijith Anand (aanand@uark.edu)	Walton College of Business (WCOB)
Key Personnel	Shakil Rafi (sarafi@uark.edu)	Department of Agriculture (AFLS)

Key Words (5):

1. Health Outcomes
2. Spatial Analysis
3. Healthcare Policy
4. Clinical Performance
5. Geographic Spillovers

Budget Amount Requested:

$30,000

Research Area(s) Addressed:

Health Outcomes, Deep Learning, Spatial Modeling, Policy Research, Clinical Decision making

Synopsis (1,000 character max):

Patient outcomes are influenced not only by their primary hospitals but also by nearby institutions due to Accountable Care Organizations (ACOs). Existing models often fail to capture the nonlinear and dynamic spillover effects of clinical outcomes across hospitals, limiting policy effectiveness. We address this gap by integrating spatial econometrics with spatio-temporal graph attention networks (ST-GATs) to model how outcomes propagate over time and space.

Using clinical outcomes data from the Centers for Medicare & Medicaid Services (CMS), we will build spatial frameworks to better understand these diffusion patterns. Our interdisciplinary team—combining expertise in health policy, econometrics, and artificial intelligence (AI)—offers both methodological rigor and practical healthcare insight.

In Year 1, we will deliver validated predictive models for Arkansas and generate actionable insights for targeted interventions. In the long-term, this prototype will support proposals to the NSF Smart Health program, the National Library of Medicine (NIH NLM), and the Patient-Centered Outcomes Research Institute (PCORI).

Introduction

Accountable Care Organizations (ACOs) (networks of clinicians, hospitals, and healthcare providers) are set up in a way such that what happens in one hospital often affects others. For example, changes in clinical practice, policy responses, or healthcare outcomes do not stay neatly within one hospital, they "spill over" to nearby institutions, both within the ACO and beyond. These spillover effects are shaped by geography (how close hospitals are to each other) and by the state or county-level systems that manage healthcare delivery. In other words, a hospital's performance is not determined in isolation, it is influenced by what is happening around it. These phenomena are formally known as *spillover effects*.

For example, if Washington Regional Hospital drastically improved its mortality rate from childbirth, outpatient centers nearby in Northwest Arkansas would see a concomitant improvement in their childbirth mortality, as would centers nearby to those centers and hence exhibit a ripple effect throughout the Northwest Arkansas region. This is valid for any such region or hospital system in the United States, and are highly amenable to modeling.

These spillovers evolve over time, especially in response to major events like new healthcare regulations (e.g., the HITECH Act [1]) or crises such as the COVID-19 pandemic [7]. To understand and quantify these ripple effects, researchers use models that track changes across both space and time, such as spatial panel models [4] and instrumental variable approaches [3]. These tools help reveal how performance in one part of the healthcare system can influence outcomes elsewhere, sometimes in unexpected ways.

However, a key limitation in extant studies and research is that these frameworks falls short when data is high-dimensional, irregularly spaced, or exhibit nonlinear interactions [11].

To address this critical gap, our proposal offers two core innovation:

1. *Collecting high-resolution data*: Unlike most studies that rely on aggregated datasets, we aim to collect and organize high-fidelity data from multiple public and private sources, that offer exceptional spatial and temporal granularity, allowing for more precise modeling of hospital performance dynamics.
2. *Cutting-edge and novel modeling techniques*: In addition to classical spatial methods (e.g., spatial autocorrelation and spatial lag models), we will implement spatio-temporal graph attention networks (ST-GATs)[10] models ideally suited to our dense dataset and capable of capturing nonlinear, directional spillover effects across both space and time. These models have seen use in traffic modeling and are ideal for spatial and temporal modeling [6] but have yet to be used in health metrics spillover, presenting a high-risk/high-reward scenario.

This is backed by four key strengths:

1. *Interdisciplinary high-skilled team*: This project integrates expertise in econometric modeling for electronic health records (**Dr. Aranyak Goswami**), policy and management science (**Dr. Abhijith Anand**) with advanced spatial analytics knowledge (**Dr. Shakil Rafi**), offering a synergistic approach to a complex systems problem.
2. *Preliminary analyses*: Preliminary analysis reveals statistically significant spatial spillovers across facilities, highlighting the need for more advanced modeling to capture both their magnitude and underlying mechanisms.
3. *Precedent and promising follow-up*: This project is strategically designed as a launchpad for competitive external funding from major agencies that support research on spatial modeling, healthcare analytics, and decision-making systems. For instance, the NIH recently funded research employing spatiotemporal modeling to assess healthcare outcomes (e.g., Project #1R01DK136515-01A1). The NSF Smart Health and Biomedical Research in the Era of Artificial Intelligence and Advanced Data Science (NSF 23-614) program actively supports projects integrating machine learning, health policy, and systems-level optimization. Additionally, organizations like the National Institute for Health Care Management (NIHCM) have funded applied research on geographic disparities and healthcare performance. Our work aligns with these priorities and is well-positioned to compete for follow-on funding from the National Institutes of Health NIH, National Science Foundation NSF, Patient-centered Outcomes Research Institute PCORI, and the Agency for Health Research and Quality AHRQ in subsequent stages.
4. *Potential for expanded recruitment:* Our research demands significant effort in data acquisition, wrangling, and management representing a technically intensive task well suited for graduate students and post-docs, contingent on the availability of substantial future funding

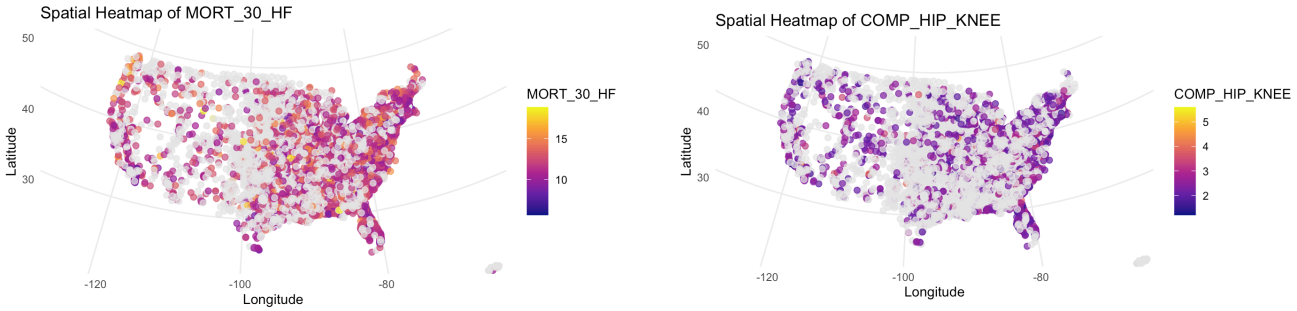
Preliminary Results

To establish proof of concept, we conducted a preliminary analysis using two representative hospital quality metrics that are reported to CMS:

1. `COMP_HIP_KNEE`: the complication rate following elective primary total hip and/or total knee arthroplasty.
2. `MORT_30_HF`: the 30-day mortality rate following heart failure hospitalization.

These metrics were selected based on their clinical relevance, availability in our dataset, and their historical use in policy evaluation and hospital benchmarking. Together, they reflect both surgical outcomes and chronic care performance, offering a robust testbed for our modeling approach.

Initial spatial analysis revealed clear patterns of regional clustering and performance diffusion, particularly among outpatient centers in proximity to major hospital hubs, see Figure 1. These patterns are informed by spillover effects known in previous literature [5], [2], [9]. Early runs of our spatio-temporal models show promise in capturing both directionality and intensity of influence between facilities—laying a strong foundation for the expanded analysis proposed in this grant.



(a) Deaths

(b) Complications

Figure 1: Side-by-side visualizations of deaths and complication metrics across hospitals

The clustering results are backed up by spatial autoregression ρ values indicating spillover effects for `MORT_30_HF` but not for `COMP_HIP_KNEE`, (see Table 1). More concretely, for every one unit increase in in `MORT_30_HF` at a flagship hospital there is an increase in `MORT_30_HF` of nearby hospitals by approximately 0.293 with a high degree of statistical certainty.

Simulating improvements in a specific hospital metric, allows us to estimate the spillover impact of targeted interventions. This enables us to identify which hospitals act as *influence hubs*, where localized improvements could lead to *system-wide gains*. For example, preliminary simulations using `MORT_30_STK` (30-day mortality rate following stroke hospitalization) suggest that a one-unit improvement in a high-impact facility can lead to average gains of up to 0.6 units in regions like Arkansas, Alabama and southern Arizona. Please see Figure 2.

Table 1: Spillover coefficient (ρ) for selected hospital metrics

Metric	$\boldsymbol{\rho}$
`MORT_30_HF`	0.293
`COMP_HIP_KNEE`	0.029

As illustrated in Figure 2, several facilities in Arkansas exhibit disproportionately high spillover scores, suggesting that targeted quality improvements in this region could yield broad regional and potentially national benefits. These results underscore the relevance of this work not just for national policy modeling, but also for state-level strategic planning in Arkansas and the broader southern U.S. healthcare system.

The Proposed Method

We will begin by collecting and integrating data from public and proprietary sources, primarily from CMS. This will require significant preprocessing and validation and is expected to take approximately one month, led by the Ph.D. student.

With our robust dataset, we will begin by estimating spatial lag models across all 143 available hospital quality metrics to quantify the direction and magnitude of spatial dependencies. Each model requires the inversion of a 4085×4085 spatial weight matrix demanding significant computational resources and careful tuning for stability and interpretability.

We will then implement a spatio-temporal graph attention network (ST-GAT), initially focusing on healthcare facilities within the state of Arkansas before scaling to a nationwide model. The ST-GAT architecture is specifically designed to capture nonlinear relationships across both space and time and has been successfully applied in domains such as traffic flow forecasting and urban mobility modeling [8]. To the best of our knowledge this has yet to be applied to model health outcomes spillover effects, presenting a high-risk/high-reward scenario.

After initial implementation, we will iteratively refine and train our models, validating their performance against classical spatio-temporal econometric models to establish robust benchmarks.

This pipeline represents a substantial technical undertaking, well-suited to the scope of a Master's thesis or PhD dissertation and offers valuable training opportunities in applied machine learning, network science, and spatial econometrics.

Outcomes and Actionable Deliverables

We propose to have the following timeline for our deliverables.

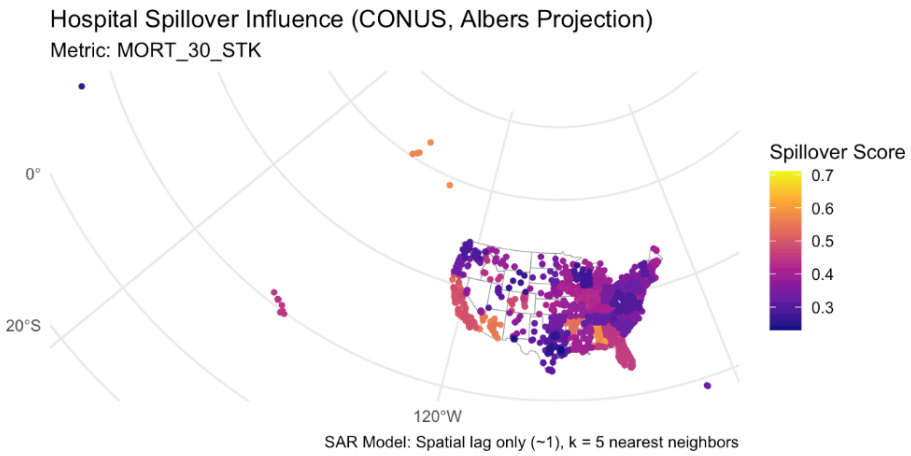


Figure 2: Hospital Spillover effects showing strong spillover in Arkansas, Alabama, and southern Arizona.

- **Month 1—2**: Data acquisition and preprocessing.
- **Month 3—4**: Spatial lag model estimation across selected hospital metrics.
- **Month 5—7**: Development and training of ST-GAT models for Arkansas facilities.
- **Month 8—9**: Nationwide model scaling and benchmarking against traditional methods.
- **Month 10—11**: Analysis of influence hubs and policy simulations.
- **Month 12**: Drafting manuscripts, submitting grant applications, and preparing the final report.

We intend to present our preliminary findings at the 2026 Joint Statistical Meetings (JSM), scheduled for August 2026 in Boston, Massachusetts. To meet this goal, we will prepare and submit a poster abstract by the December 2025 submission deadline. Concurrently, we aim to make a preprint of our analysis publicly available on `arXiv` by Spring 2026, accompanied by open-source code and data repositories hosted on GitHub to promote transparency and reproducibility. For final publications, we will target high-impact peer-reviewed journals such as *Health Economics*, and *Econometrica*.

Budget Justification

This budget strategically supports the successful execution and dissemination of the proposed research over a 12-month period. Part of a Ph.D. student's salary will be supported at $1,800 per month for three months, providing critical expertise in data modeling, machine learning implementation, and manuscript development. Travel costs ($4,000) will support attendance at a domestic conference to present findings and foster collaborations. Materials and supplies ($7,500) cover data storage and collaboration tools necessary for secure handling and version control of high-resolution healthcare datasets.

Publication fees ($1,602) are allocated for two open-access journal articles to ensure broad visibility and rapid dissemination of results. Other direct costs ($6,500) are designated for cloud-based GPU access (e.g., AWS/GCP credits) and essential licensed software (e.g., MATLAB toolboxes, GraphML add-ons) required for implementing advanced spatiotemporal models. No institutional overhead is requested, ensuring that the entire budget directly supports research, dissemination, and project infrastructure.

Conclusion

This project delivers a novel, interdisciplinary framework to model how healthcare outcomes propagate through institutional networks over time and space. By integrating high-resolution hospital data with cutting-edge spatiotemporal machine learning, we address a major gap in understanding systemic performance spillovers, a challenge central to health systems optimization.

Preliminary results reveal strong influence patterns and identify high-impact facilities in Arkansas where localized improvements may yield broader regional gains. These findings have immediate relevance for hospital operations, health policy, and strategic planning across disciplines.

In one year, we will produce validated predictive models, identify actionable intervention hubs, and publicly share open-source tools, preprints, and publications. These outputs represent a strong foundation for proof-of-concept success and future external funding. The project exemplifies the type of high-risk/high-reward, cross-college collaboration this seed grant seeks to support. It brings together expertise from business, engineering, health data science and with clear alignment to external programs like NSF Smart Health, NIH, NLM, and PCORI, this initiative positions the University of Arkansas for leadership in scalable, data-driven healthcare research.

References

[1] THE HITECH ACT—an overview. *AMA Journal of Ethics*, 13(3):172–175.

[2] Badi H. Baltagi and Yin-Fang Yen. Hospital treatment rates and spillover effects: Does ownership matter? *Regional Science and Urban Economics*, 49:193–202.

[3] Andrew Chesher and Adam M Rosen. What do instrumental variable models deliver with discrete dependent variables? *American Economic Review*, 103(3):557–562.

[4] J. Paul Elhorst. Spatial panel data models. In Manfred M. Fischer and Arthur Getis, editors, *Handbook of Applied Spatial Analysis*, pages 377–407. Springer Berlin Heidelberg.

[5] Igor Francetic, Rachel Meacock, Jack Elliott, Sren R. Kristensen, Phillip Britteon, David G. Lugo-Palacios, Paul Wilson, and Matt Sutton. Framework for identification and measurement of spillover effects in policy implementation: intended non-intended targeted non-targeted spillovers (INTENTS). *Implement Sci Commun*, 3(1):30.

[6] Xiangyuan Kong, Weiwei Xing, Xiang Wei, Peng Bao, Jian Zhang, and Wei Lu. Stgat: Spatial-temporal graph attention networks for traffic flow forecasting. *IEEE Access*, 8:134363–134372, 2020.

[7] Yalini Senathirajah, David R Kaufman, Kenrick Cato, Kenrick Cato, Pia Daniel, Bonnie Arquilla, Patricia Roblin, Andre Kushniruk, Elizabeth M. Borycki, Emanuel Feld, and Poli Debi. The impact of COVID-19 regulatory reporting burden: A comparative study of a small independent hospital and a large network hospital (preprint). *Online Journal of Public Health Informatics*.

[8] Junho Song, Jiwon Son, Dong-hyuk Seo, Kyungsik Han, Namhyuk Kim, and Sang-Wook Kim. St-gat: A spatio-temporal graph attention network for accurate traffic speed prediction. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, CIKM '22, page 45004504, New York, NY, USA, 2022. Association for Computing Machinery.

[9] Olga Yakusheva. Health spillovers among hospital patients: Evidence from roommate assignments. *American Journal of Health Economics*, 3(1):76–107.

[10] Chenhan Zhang, James J. Q. Yu, and Yi Liu. Spatial-temporal graph attention networks: A deep learning approach for traffic forecasting. *IEEE Access*, 7:166246–166256.

[11] Xiaoyu Zhou. Essays on estimation for nonlinear spatial models.

BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. DO NOT EXCEED FIVE PAGES.

NAME: **ARANYAK GOSWAMI**

eRA COMMONS USER NAME (credential, e.g., agency login): GOSWAMIA

POSITION TITLE: Tenure Track Assistant Professor in Bioinformatics and Computational Biology, Department of Animal Sciences, University of Arkansas

EDUCATION/TRAINING (Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)

INSTITUTION AND LOCATION	DEGREE (if applicable)	Completion Date MM/YYYY	FIELD OF STUDY
University of Calcutta	MS	05/2007	Genetics Bioinformatics, Computational Biology, and Genetics
Indian Institute of Chemical Biology	PhD	07/2018	
Bose Institute, Kolkata	National Post-Doctoral Fellow	11/2019	Bioinformatics, Computational Biology, and Genetics
Yale University	Post-Doctoral Associate	04/2022	Bioinformatics, Computational Biology, and Genetics
Stanford University	Post-Doctoral Scholar	03/2024	Bioinformatics, Computational Biology, and Genetics

A. Personal Statement

I am an Assistant Professor in the Department of Animal Sciences at the University of Arkansas, specializing in computational biology and bioinformatics, with a focus on integrating host genetics and microbiome data. My Ph.D. in Bioinformatics and Microbial Genomics, along with postdoctoral research at Yale and Bose Institute, has equipped me with expertise in GWAS, microbial genetics, and advanced computational techniques such as Variational Autoencoders (VAEs) and machine learning models. Currently, my research involves analyzing large genomic datasets to uncover genetic and microbial interactions that contribute to psychiatric disorders. My experience with advanced computational methods, causal inference techniques, and high-dimensional data analysis provides a solid foundation for identifying novel biomarkers and therapeutic targets for neuropsychiatric conditions.

B. Positions, Scientific Appointments, and Honors

Positions and Employment

- 2024–present: Tenure Track Assistant Professor, Bioinformatics and Computational Biology, Department of Animal Sciences, University of Arkansas

Honors:

- 2005: UGC Junior Research Fellowship, Govt of India
- 2007: Senior Research Fellowship, Govt of India
- 2018: National Postdoctoral Fellowship, Govt of India

Scientific Memberships:

- 2019–2020: Psychiatric Genomics Consortium
- 2022–2024: American Society of Gene and Cell Therapy
- 2024–2025: American Society for Microbiology

C. Contributions to Science

1) Saha, S. K., Goswami, A., & Dutta, C. (2014). Association of purine asymmetry, strand-biased gene distribution and PolC within Firmicutes and beyond a new appraisal. BMC Genomics, 15(1), 1-26. **Impact Factor- 4.558**

2) Wendt, F. R., Pathak, G. A., Tylee, D. S., Goswami, A., & Polimanti, R. (2020). Heterogeneity and Polygenicity in Psychiatric Disorders: A Genome-Wide Perspective. Chronic Stress, 4, 2470547020924844. **Impact Factor- 4.18**

3) Wendt, F. R., Pathak, G. A., Levey, D. F., Nuñez, Y. Z., Overstreet, C., Tyrrell, C., ... & Polimanti, R. (2021). Sex-stratified gene-by-environment genome-wide interaction study of trauma, posttraumatic stress, and suicidality. Neurobiology of stress, 14, 100309. **Impact Factor- 7.142**

4) Goswami, A., Chowdhury, A. R., Sarkar, M., Saha, S. K., Paul, S., & Dutta, C. (2015). Strand-biased gene distribution, purine asymmetry and environmental factors influence protein evolution in Bacillus. FEBS letters, 589(5), 629-638. **Impact Factor- 3.864**

5) Pathak, G. A., Wendt, F. R., De Lillo, A., Nunez, Y. Z., Goswami, A., De Angelis, F., ... & Polimanti, R. (2021). Epigenomic Profiles of African American Transthyretin Val122Ile Carriers Reveals Putatively Dysregulated Amyloid Mechanisms. Circulation: Genomic and Precision Medicine, 14(1), e003011. **Impact Factor-7.465**

6) Wendt, F. R., Pathak, G. A., Levey, D. F., Nuñez, Y. Z., Overstreet, C., Tyrrell, C., ... & Polimanti, R. (2020). Trauma and posttraumatic stress interact with sex-specific risk loci for suicidality and converge on brain extracellular matrix biology and synaptic plasticity. medRxiv.

7) Goswami, A., Wendt, F. R., Pathak, G. A., Tylee, D. S., De Angelis, F., De Lillo, A., & Polimanti, R. (2021). Role of microbes in the pathogenesis of neuropsychiatric disorders. Frontiers in Neuroendocrinology, 62, 100917. **Impact Factor-8.333**

8) De Angelis, F., Wendt, F. R., Pathak, G. A., Tylee, D. S., Goswami, A., Gelernter, J., & Polimanti, R. (2021). Drinking and smoking polygenic risk is associated with childhood and early-adulthood psychiatric and behavioral traits independently of substance use and psychiatric genetic risk. Translational psychiatry, 11(1), 1-12. **Impact Factor- 7.989**

9) Pathak, G. A., Wendt, F. R., Goswami, A., Koller, D., De Angelis, F., Polimanti, R., & COVID-19 Host Genetics Initiative. (2021). ACE2 Netlas: In silico Functional Characterization and Drug-Gene Interactions of ACE2 Gene Network to Understand Its Potential Involvement in COVID-19 Susceptibility. Frontiers in genetics, 1523. **Impact Factor- 4.772**

10) COVID-19 COMPLICATIONS AND SUGGESTED MEASURES: MODERN TOOLS FOR INTERVENING PANDEMIC JJ Yusuf, K. Z., Ansar, W., Goswami, A., Mandal, S., Tahrim, H., Poddar, S ...
Journal of Health and Translational Medicine, 25 (1), 145-153

11) Pekrun, K., Stephens, C. J., Gonzalez-Sandoval, A., Goswami, A., Zhang, F., Tarantal, A. F., ... & Kay, M. A. (2024). Correlation of antigen expression with epigenetic modifications after rAAV delivery of a hyperactive human Factor IX variant in mice and rhesus macaques. Molecular Therapy. **Impact Factor- 12.4**

12) Puzzo, F., Crossley, M. P., Goswami, A., Zhang, F., Pekrun, K., Garzon, J. L., ... & Kay, M. A. (2024). AAV-mediated genome editing is influenced by the formation of R-loops. Molecular Therapy)
Impact Factor- 12.4

[* I have been involved in all relevant analysis, writing, and designing of figures for all my First Author papers and I have been involved in the conception, writing, and part of analytic execution in my second author papers.]

BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. **DO NOT EXCEED FIVE PAGES.**

NAME: ***ABHIJITH ANAND***

eRA COMMONS USER NAME (credential, e.g., agency login): -NA-

POSITION TITLE: ***Assistant Professor of Information Systems, Department of Information Systems, Walton College of Business, University of Arkansas***

EDUCATION/TRAINING

INSTITUTION AND LOCATION	DEGREE	Completion Date	FIELD OF STUDY
KSIT, Visvesvaraya Technological University Bangalore, India	Bachelors	05/2009	Engineering, Electronics and Communications
University of Wollongong (UoW) Wollongong, Australia	Masters	10/2011	Information and Communication Technology
University of Wollongong Wollongong, Australia	Masters – Research	06/2013	Information Systems and Technology
University of Waikato, Waikato School of Management, New Zealand *Note: Parts of my PhD research were conducted at the University of Wollongong, McGill University and University of Technology Sydney*	PhD	11/2017	Information Systems
University of Wollongong Wollongong, Australia *Note: I have held multiple positions during my PhD, hence dates overlap.*	Global Challenges Scholar	04/2016	Engineering and Information Sciences
University of Wollongong Wollongong, Australia *Note: I have held multiple positions during my PhD, hence dates overlap.*	Associate Research Fellow	08/2017	Management Science

A. Personal Statement

I am an Assistant Professor in the Department of Information Systems at the Walton College, University of Arkansas. My research focuses on the intersection of IS value, healthcare information systems, IS strategy, and mathematical modeling. A core theme of my work is understanding how artificial intelligence and data analytics transform organizational strategies and competitive actions, with a particular emphasis on healthcare institutions.

My PhD in Information Systems, along with fellowships and roles on several externally funded grants in Australia, has equipped me with deep expertise in healthcare IT, analytics, and AI. Methodologically, I employ a range of advanced techniques—including statistical and econometric analysis, causal inference, computational modeling, machine learning, and mathematical modeling tailored to theoretical testing.

One stream of my research work examines the behaviors between healthcare practitioners and healthcare analytics platforms – aligned with the key objective of the HITECH Act. A significant strand of my research investigates how hospitals use analytics systems to respond to performance pressures in the complex and resource-constrained clinical environment. This work addresses a critical gap in our understanding of when and why healthcare practitioners turn to analytics to guide clinical decision-making.

Another important strand of my research focuses on the role of contextual factors in shaping the clinical and public health interpretation of vaccine efficacy, particularly in high-stakes pandemic environments. My work reveals that geographic regions with higher disease prevalence during clinical trials tend to report lower vaccine

efficacy - highlighting the need for more nuanced, context-aware approaches to evaluating and communicating vaccine outcomes. This has implications for both scientific understanding and public policy during pandemics.

Lastly, another area of my research focuses on how organizations and societies can derive value from emerging technologies like analytics and AI. I've developed frameworks such as the Theory of Performance-Driven Search, a two-stage analytics investment model, and the Performance Attainment Index to support data-driven decision-making. My work also offers guidance for prioritizing AI capabilities based on organizational needs and temporal context. Extending beyond firms, I study how national technology networks - through connectivity, institutional support, and local innovation - enable poverty reduction and disaster resilience. This research strand highlights the broader societal value of digital investments, particularly in improving equity and public health.

My research has also had significant industry reach. I have presented findings at practitioner-oriented forums such as SIM Connect, MIS Quarterly Executive, SAS Inc., Westpac, Ports Australia, and the Australian Tax Office. My work has informed whitepapers, business reports, and industry blogs, and has been featured in outlets including Science Magazine, Technology Networks, EurekAlert!, Medical Xpress, and Newswise. Notably, SAS Inc. incorporated insights from my research into their corporate training materials, with a public-facing report receiving over 12,000 downloads - demonstrating strong practitioner engagement and real-world impact.

In summary, I bring the expertise, leadership experience, training, and motivation not only to successfully collaborate on and execute the proposed research projects, but also to generate high-impact publications suitable for top-tier academic journals and actionable outputs for practitioners.

B. Positions, Scientific Appointments, and Honors

Jan 2018 – Present: Assistant Professor, Department of Information Systems, Walton College of Business, University of Arkansas
Aug 2017 – Dec 2017, Research Instructor, Department of Information Systems, Walton College of Business, University of Arkansas

Associate Editor Appointments

1. CTO Division, Academy of Management Annual Meeting *(**AOM**)* 2025, *Copenhagen, Denmark*
2. Digital Strategy, *International Conference on Information Systems (**ICIS**) 2024, Bangkok, Thailand*
3. Special Issue on Digital Organization (2023), Information Systems Frontier *(**ISF**)*
4. Governance, Digital Strategy, and Value, ***ICIS** 2023, Hyderabad, India*
5. HCI and Robotics, *Pacific Asia Conference on Information Systems (**PACIS**) 2023, Nanchang, China*
6. General IS Topics, ***ICIS** 2022, Copenhagen, Denmark*
7. General IS Topics, ***ICIS** 2021, Austin, USA*
8. IT Strategy, Leadership and Governance, ***PACIS** 2021, Dubai, UAE.*
9. Special Issue on Interpretable AI-enabled Online Behavior Analytics (2020), *Internet Research* Journal.
10. General IS Topics, *International Conference on Information Systems (**ICIS**) 2020, Hyderabad, India.*
11. Advances in Research Methods, ***ICIS** 2020, Hyderabad, India*
12. Transforming Society with Digital Innovation Track, ***PACIS** 2019, Xi'an, China*

Honors, Grants and Fellowships (Key Highlights): EJIS Outstanding Reviewer Contributions, Pandemic Research Recovery Grant, AIS/ACM Best Doctoral Dissertation Award in the Field, PHIS-NZ Best Doctoral Dissertation Award for IS Research New Zealand (2018), Ports Australia Fellowship, SAS Inc. Australia Research Fellowship, ICIS Doctorial Consortium Fellow, Global Challenges Fellowship and among others.

Scientific Memberships: 2014 – Present: Association of Information Systems, 2022 – Present: Informs

C. Contributions to Science: Please see:
https://scholar.google.com/citations?user=4QxVNC8AAAAJ&hl=en

BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. **DO NOT EXCEED FIVE PAGES.**

NAME: **SHAKIL RAFI**

eRA COMMONS USER NAME (credential, e.g., agency login): SARAFI

POSITION TITLE: POST-DOCTORAL FELLOW

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	Start Date MM/YYYY	Completion Date MM/YYYY	FIELD OF STUDY
TROY UNIVERSITY, TROY, AL	BSC	01/2011	05/2015	PURE MATHEMATICS, PHILOSOPHY
UNIVERSITY OF ARKANSAS, FAYETTEVILLE, AR	MSC	08/2016	12/2019	PURE MATHEMATICS
UNIVERSITY OF ARKANSAS, FAYETTEVILLE, AR	PHD	01/2020	05/2024	APPLIED MATHEMATICS
UNIVERSITY OF ARKANSAS, DALE BUMPERS COLLEGE OF AGRICULTURAL, FOOD, AND LIFE SCIENCES, FAYETTEVILLE, AR	POST-DOCTORAL FELLOW	08/2024	PRESENT	BIOINFORMATICS AND COMPUTATIONAL BIOLOGY

A. Personal Statement

I am a Post-Doctoral Fellow at the Department of Agriculture at the University of Arkansas. My PhD was in Applied Mathematics where I focused on developing an algebraic framework for understanding feed-forward neural networks and gave strict upper bounds on the complexity of these neural networks in order to approximate certain solutions to differential equations. I am therefore also an expert in advanced machine learning techniques such as variational autoencoders and attentive transformers. In addition to my theoretical underpinnings, I have extensive knowledge and experience applying machine learning to real-world problems, for example in modeling the maternity cycle of clients at Arkansas Blue Cross and Blue Shield, investigating the inequality of ride-sharing access in Chicago during the pandemic, and revealing discrepancies in loan access during the pandemic of women and black owned businesses during the pandemic. My strong background in mathematics, and my years of experience translating theoretical concepts to real-world applications make me an ideal computational lead for this project.

B. Positions, Scientific Appointments and Honors

Positions and Employment
2024-Present Post-doctoral Fellow, Dale Bumpers College of Agriculture.
2023-2024 Lecturer, Department of Data Science, Sam M. Walton College of Business, University of Arkansas.
2023-2023 Intern, Health Economics Team, Arkansas Blue Cross and Blue Shield
2015-2016 Asst. Business Analyst. R.S.S. Wears Ltd.

Scientific Memberships
2022-2024 Treasurer, Graduate Student Colloquium, Department of Mathematics
2023, 2024 Lead Judge, Computer Science Category, Senior, Northwest Arkansas Regional Science and Engineering Fair.
2016-Present American Mathematical Society
2022-Present Society for Industrial and Applied Mathematics
2014-Present Alabama Eta Chapter of Pi Mu Epsilon, the national mathematics honor society.

Honors
2024: The SIAM travel grants award, Fayetteville, AR
2023: Selected by my department to attend the Summer Graduate School in Machine Learning at the Simons Laufer Mathematical Sciences Institute (formerly Mathematical Sciences Research Institute) at UC San Diego. San Diego, CA.
2022: Privileged to attend the Arkansas Summer Research Institute in Machine Learning, an NSF EPSCOR funded project. Fayetteville, AR
2021: The Lawrence Jesser Toll, Jr. Endowed Fellowship, Fayetteville, AR
2017: The Bangladesh-Sweden Travel Fund, Dhaka, Bangladesh
2011: The Millenium Scholar's Award, Troy University, Troy, AL

C. Contributions to Science

Rafi, S. *Gender Disparities in Arkansas, and Income Disparities in the US for the PPP Loan Program*. Preprints 2023, 2023090654. [Preprint]. Available from: https://doi.org/10.20944/preprints202309.0654.v1

Shakil Rafi , Arna Nishita Nithila . *Who rides Uber anyway? A census-tract level analysis and clustering of ride-shares for the city of Chicago during the era of the pandemic.* [Preprint]. TechRxiv. 2022. DOI: 10.36227/techrxiv.21076042.v2

Shakil Rafi. *A Clustering Look at Chicago Rideshares.* Poster session presented at: SIAM Mathematics of Data Science Conference '22. 2022. San Diego, CA.

Rafi S, Padgett J (2024). *nnR: Neural Networks Made Algebraic*. R package version 0.1.0, <https://CRAN.R-project.org/package=nnR>.

Rafi S., Padgett J., Nakarmi U. *Towards an Algebraic Framework For Approximating Functions Using Neural Network Polynomials*. arXiv. 2024. DOI: https://doi.org/10.48550/arXiv.2402.01058

Rafi S. *Maximal Parameter Estimates for Neural Networks and Uncertainties in Approximation*. Poster session presented at: SIAM Conference on Uncertainty Quantification '24. Trieste, Italy.

Rafi S. *Analysis and Construction of Artifcial Neural Networks for the Heat Equations, and Their Associated Parameters, Depths, and Accuracies.* [dissertation]; University of Arkansas. 2024, 303p.

Full list of publications and preprints can be found in Google Scholar:
https://scholar.google.com/citations?user=xbb3YbIAAAAJ&hl=en

BUDGET - University of Arkansas		Date:									
Proposed to (Sponsor):	Collaborative Research Program										
Proposed Start & End Dates:	2025-07-01	to	2026-06-30					Select project type:	On-campus research		
UA Lead Investigator:	Dr. Aranyak Goswami										
UA Lead Dept/College:	Department of Agriculture										

		Type		Person-Months		Cost	Year 1			Total
SALARIES & WAGES	Base Salary	Appointment		CAL/AY	SMR	Share	PI 1	PI 2	PI 3 (co-PI)	UA
Postdoctoral Associate		12 mo.	NonCL			0%	0	0	0	0
Research Associate (staff)		12 mo.	NonCL			0%	0	0	0	0
Research Assistant or Tech		12 mo.	NonCL			0%	0	0	0	0
Graduate Assistant (Ph.D.)		1	3 mo. @	1800		0%	5,400	0	0	5,400
Graduate Assistant (Masters)			mo. @			0%	0	0	0	0
Hourly, non-student(s)			hrs @			0%	0	0	0	0
Hourly, enrolled student			hrs @			0%	0	0	0	0
Total S&W							*5,400*	*0*	*0*	5,400
FRINGE BENEFITS			Institutional Rate:							
GA salary			5.70%				308	0	0	308
Hourly wages			6.70%				0	0	0	0
Enrolled student wages			0.70%				0	0	0	0
Total FB							*308*	*0*	*0*	308
Total Salaries + Benefits							**5,708**	**0**	**0**	**5,708**
TRAVEL - Domestic							2,000		2,000	4,000
MATERIALS & SUPPLIES (not *fees or services, which are "Other"*)							7,500			7,500
JOURNAL PUBLICATION FEES							1,602			1,602
OTHER DIRECT COSTS *(Itemize by type; insert extra rows if needed.)*										
GPU Hosting, Storage, and Service Fees							6,500		6,500	0
									0	0
									0	0
									0	0
									0	0
									0	0
Subtotal Other Direct Costs							**6,500**	**0**	**6,500**	**0**
Modified Total Direct Costs (above subtotal costs subject to F&A Cost)							*23,310*	*0*	*23,310*	*0*
F & A COST (MTDC x RATE):		**0.0%**					**0**		**0**	
F & A COST (UNRECOVERED):		**0.0%**						**0**		**0**
F & A COST (COST-SHARE):		**0.0%**						**0**		**0**
Modified Total Direct Costs (first $25K of each subaward)							*0*		*0*	
F & A COST (MTDC x RATE)SUB(S):		**0.0%**					**0**		**0**	
(Direct Costs not subject to F&A Cost, with the exception that the first $25K of each subaward is subject to F&A):										
GRA TUITION	# Credit Hours: 15			Rate: $446			6,690		6,690	0
PARTICIPANT (TRAINEE) STIPEND									0	0
PARTICIPANT (TRAINEE) TRAVEL									0	0
PARTICIPANT (TRAINEE) SUBSISTENCE									0	0
PARTICIPANT (TRAINEE) OTHER									0	0
SUBAWARD #1, total	(Institution):								0	0
SUBAWARD #2, total	(Institution):								0	0
SUBAWARD #3, total	(Institution):								0	0
SUBAWARD #4, total	(Institution):								0	0
TOTAL DIRECT COST							**30,000**	**0**	**30,000**	**0**
TOTAL PROJECT COST							**$30,000**	**$0**	**$30,000**	**$0**