

SHUN ZHANG

QUALIFICATIONS PROFILE

- Proven leader in building and managing high-performing technical teams, with expertise in generative AI, advanced data science, and cloud platforms.
- Over 15 years of experience driving AI innovation and delivering impactful solutions.
- Strong track record of aligning data-driven initiatives with business goals, managing complex stakeholder relationships, and leading teams to execute strategic objectives

PROFESSIONAL EXPERIENCE

ASTRAZENECA, SOUTH SAN FRANCISCO, CA

Director of AI and Data

START ON OCT 2023 - CURRENT

- **AZ Brain Application Development and Delivery:** Led the development and delivery of machine learning solutions in partnership with senior leaders across Insight & Analytics, Engineering, and Market Research. These solutions support Oncology and Biopharma brands by driving strategic priorities, including patient enrichment, dynamic targeting, and network analysis. Enabled the optimization of Key Opinion Leaders (KOLs) and healthcare provider identification to enhance commercial strategies and marketing solutions.
- **Generative AI Application:** Developed the first in-production enterprise-level generative AI solution, leveraging large language models to analyze patient journeys, extract insights from claims data and medical records, and deliver actionable intelligence to enterprise and field users. Compared foundational models such as GPT-4, Claude-3.5, and open-source frameworks to accelerate AI-driven solutions for Medical, Commercial, and R&D teams. Designed safeguards against hallucinations to ensure AI governance, compliance, and regulatory adherence. Built a scalable, reusable infrastructure to support global integration and expansion of AI solutions.
- **MLOps / LLMOps Application Development:** Led the creation of MLOps applications to provide an end-to-end platform for managing and deploying data science models, ensuring transparency and communication between technical teams and business users. Streamlined model transitions from development to production, with a focus on scalability, reliability, and business alignment. Implemented continuous monitoring and adaptation processes to generate insights, supporting data-driven decision-making. Expanded AI and machine learning accessibility for diverse stakeholders to improve transparency and business effectiveness.
- **Team Leadership and Organizational Transformation:** Built and led a high-performing technical team, streamlining analytical processes and workflows. Provided mentorship to managers and nurtured the growth of junior members, fostering a collaborative and learning-focused environment. Spearheaded organizational transformation by championing innovation and a culture of experimentation, driving the successful adoption of cutting-edge capabilities across the company.
- **Enterprise AI Solution Initiative:** Led the initiative to design and deploy enterprise AI solutions, emphasizing scalability, security, and regulatory compliance. Key components include the integration of advanced machine learning models with cloud platforms (AWS, Azure) and enterprise systems via robust APIs. Developed a modular, reusable infrastructure to support AI model lifecycle management, including automated deployment, version control, and continuous monitoring for performance and model drift. Established comprehensive data governance frameworks and ensured adherence to AI governance and compliance standards. Enhanced operational efficiency through the adoption of advanced analytics, dynamic automation, and AI-driven insights across business functions, driving enterprise-wide digital transformation and decision-making capabilities.

ALIBABA CLOUD BELLEVUE, WA

Sr. Director of Health and Life Science R&D

SEPT 2021 – JUN 2023 (DUE TO COMPANY LAYOFF)

- **12 million population City Brain data platform:** led the successful implementation of a flagship data platform for Hangzhou City Brain, serving a population of 12 million individuals. My responsibilities included end-to-end management of the project, from requirements gathering and data modeling to ETL development, data warehousing, and data visualization. Leveraging a variety of platforms, including Alibaba Cloud, Oracle, SQL Server, and Hadoop, I ensured that the platform was optimized for performance, scalability, and data accuracy. My expertise in project and stakeholder management proved invaluable in coordinating multiple teams and ensuring that project milestones were achieved on time and within budget. The platform has since been used to develop multiple mobile phone applications, including online health data sharing, patient triage, chronic health management, and population health management. My ability to deliver data-driven insights that align with business objectives has been instrumental in driving innovation and improving outcomes.
- **Pharmaceutical company (Roche) enterprise data platform:** built and managed data warehousing and data lake solutions for enterprise data storage, retrieval, and analysis of large datasets. I developed ETL processes to meet specific business requirements and utilized data integration platforms such as MuleSoft and Alibaba ESB to integrate

disparate systems, applications, and data sources into a unified view of data. I have built data quality tools to ensure data accuracy, completeness, and consistency. preparing white papers for data documentation and knowledge sharing with internal and external teams.

- **Large Language Model application in Healthcare:** Employing advanced Qwen (LLM foundational model developed by Alibaba Cloud) to create healthcare solutions to enhance medical coding processes, produce comprehensive patient overviews, and supply pertinent medical knowledge suggestions. By precisely analyzing and decoding medical information, these tools boost efficiency and precision in the medical coding domain. The application concise patient summaries containing vital medical details, which contribute to better diagnostic and therapeutic outcomes. By offering relevant medical knowledge and guidance to healthcare professionals, these applications foster a deeper comprehension of intricate medical concepts, ultimately resulting in superior patient care. Incorporate the LLMs model into the Clinical Decision Support System (CDSS) and integrate this module into the hospital's Electronic Health Record (EHR) system to assist in primary care diagnosis and treatment procedures.
- **AI for Drug Discovery High Performance Computation Platform:** Built a public cloud portal and GPU-supported protein function/structure prediction tool for AI-driven drug discovery, based on Alibaba Max Compute and Elastic Cloud Services (ECS). Developed a next-generation R&D high-performance computing (HPC) platform with full-stack capabilities, from secure IDC to stable and robust infrastructure, AI scenario-optimized platform, integrated GPU and CPU computing, and orchestration of on-premises and elastic cloud resources for innovation.
- **Medical Knowledge Engineer Project:** Initiated a healthcare knowledge graph platform to store, extract, manage, and update healthcare-related knowledge. Utilized Transformer architecture to automate healthcare knowledge extraction from speech, documents, and clinical guidelines to support healthcare GenAI applications.
- Ensured compliance (**HIPPA/GDPR/China data security law** etc.) and security requirements were met while providing support to operations and scientists.
- Led strategy and solution discussions to accelerate digital transformation across payers, providers, and pharma companies. Provided thought leadership focusing on improving customer experiences through business solutions incorporating machine learning, AI, cloud infrastructure, and APIs.
- Managed a team of over 45 data scientists, data engineers, software engineers, product teams, and vendors to deliver high-quality AI products in healthcare and life science settings.

MICROSOFT REDMOND, WA

Director, AI for Health

JAN 2020– SEPT 2021

- Built Azure cloud platforms for hospitals, universities, and non-profit organizations.
- Provided legal consultation, compliance review, and contractor negotiation for all Microsoft AI for Health projects.
- Led teams in data architecture, informatics, data engineering, and data science to support research, clinical care, and open-source machine learning algorithms.
- Managed external partner relationships.
- Provided internal leadership for data analysis reporting.
- Key project includes:
 - Contributed to Microsoft's COVID-19 global response analysis and data initiative.
 - Participated in the COVID-19 High Performance Computing Consortium initiative.
 - Developed a machine learning model to detect prostate cancer based on PET/CT images for the British Columbia Cancer Research Center.
 - Contributed to the Novartis Foundation AI for Leprosy and AI for Better Heart projects by evaluating the risk of patient leprosy using leprosy images and build model for cardiovascular disease risk stratification.
 - Built a genetic database research and data sharing portal for sudden infant death syndrome at Seattle Children's Hospital.
 - Contributed to the Nationwide Children's AI-enhanced motion captured video project to predict infant brain development.
 - Worked with SRL Diagnostics to enhance pathology diagnosis with AI.

HEALTH CARE SERVICES CORPORATION (BLUE CROSS BLUE SHIELD) CHICAGO, IL

Sr Director of Data Science

JUL 2016 – DEC 2019

Provider Quality Evolution

- Designed a provider quality measure matrix based on NCQA, HEDIS, and clinical guidelines.
- Managed and processed insurance claims data and electronic health records data to support enterprise quality data requests and analyses.
- Planned, managed, and controlled the activities of enterprise quality measurement that provides business intelligence support for clinical operations.
- Conducted A/B testing and project evaluations.
- Evaluated and assessed the Medicare Advantage Quality Improvement and STARS rating project.

- Led monthly, quarterly, and annual budgeting and forecasting processes, ensuring strategic alignment across senior leadership.
- Assisted in the preparation of leadership presentations, financial reporting, and industry/peer group benchmarking.
- Provided population health management data support using internal data, census data, HRSA data, and vendor data.

Provider and Network Decision Analysis (PANDA)

Drive PANDA business requirements to meet diverse customer needs

- Create patient-level risk adjustment models for patients by utilizing machine learning model and leveraging claims data, external socioeconomic data, and comorbidity data.
- Regroup diseases with similar treatment and cost patterns for physician comparison by utilizing clustering analysis.
- Predict physician's working specialty to correct physician's register records by utilizing ensemble models based on historical claims and patient volume.
- Leverage machine learning models to predict the efficacy of providers who don't have enough claims data but exist in the insurance network.

Interface with other business teams and projects/programs to expand PANDA capabilities and impact internally and externally to providers, network teams, and strategy teams.

Develop PANDA reporting capabilities with static and dynamic visualization and detailed data attributions to meet the highest professional standard on both usability and usefulness.

Provider Social Network Visualization Tool Development

Developed an interactive visualization tool for internal users to contract with providers.

- Optimized external providers' patient referral patterns by analyzing referral data and providing insights for improvement.
- Created social network visualizations using Python Bokeh to help providers and network teams understand the relationships among different physicians and their referral patterns.

NORC AT THE UNIVERSITY OF CHICAGO CHICAGO, IL

Statistician III

Nov 2014 – JUL 2016

- ***U.S. Department of Health and Human Services & OPTUM - Rheumatoid Arthritis Data Challenge (2016 Health Datapalooza):***

Wrote a proposal to attend the 2016 Health Datapalooza Data Challenge and won first prize at the final competition. Used integrated EMR, claims data, and free text provider notes data to build a predictive model to estimate Rheumatoid Arthritis patients' medication persistence. Reported findings and solutions to the committee.

- ***Medicare and Medicaid Beneficiary Survey***

Evaluated disclosure risk for releasing public use data. Linked claims data with survey data to perform analytical reports for Medicare Fee for Service patients. Managed data delivery and data quality control for multi-round data collections.

- ***Health Care Cost Institute Data Enclave Project***

Managed billions of private health insurance claims data from major health insurance companies in the US. Explained the variation in physician payment rates within and across markets. Compared different providers' health efficiency and costs between different markets.

- ***Multiple System Utilization for Tricare Beneficiaries Project***

Managed a longitudinal claims database for 9.6 million active-duty personnel, military retirees, and dependents. Examined and monitored migration behavior and payment between Tricare beneficiaries between the direct care and purchased care system. Compared the quality of care, measured through processes of care and outcome of care in military facilities and civilian providers.

Used the Medical Episode Group (MEG) developed by Thomson Reuters to compare payments for care.

NATIONAL CENTER FOR PRIMARY CARE ATLANTA, GA

Biostatistician

Jun 2011 – Nov 2014

- ***Social Network Analysis: Breast Cancer Patient-Provider Network***

Conducted social network analysis on the breast cancer patient-provider network to measure social network properties and identify important nodes in the healthcare network. Compared communities with different network properties to explain differences in health outcomes.

- ***Medical Model Simulation and Decision-making Project***

Developed a systematic framework for medical decision-making using TreeAge software. Created Markov models, Monte Carlo simulations, and decision tree analyses to evaluate available options and develop strategies that maximize success while minimizing risk. Conducted cost-effectiveness analysis for adherence to long-term controller medications in children with asthma.

CENTER FOR DISEASE CONTROL AND PREVENTION, SHANGHAI(JINGAN), CHINA

Assistant of Director

Jul 2007 – Aug 2009

•HIV-TB Global Fund and Gates Foundation Initiative

Led the operational management of a high-impact project in partnership with the Gates Foundation and Global Fund, aimed at controlling HIV and tuberculosis in China. Oversaw project execution, including data collection, analysis, and reporting, ensuring that the project met the funders' requirements and goals. Streamlined operational workflows to maximize efficiency and ensured timely and accurate reporting of key insights.

EDUCATION & CREDENTIALS

- MASSACHUSETTS INSTITUTE OF TECHNOLOGY ○ Boston, MA (2014)
Short Courses: Data & Models in Engineering, Science, & Business
- UNIVERSITY OF NORTH TEXAS HEALTH SCIENCE CENTER ○ Fort Worth, TX (2011)
Master of Public Health with concentration in Biostatistics
- ANHUI MEDICAL UNIVERSITY ○ Hefei, Anhui, PR.China (2007)
Preventive Medicine

Awards & Recognitions

2017-2018: CIO magazine Digital Edge 50 Award in US (PANDA asset owner)

2016-2017: Health care services corporation outstanding employee award

2016(May): First Prize winner of 2016 Health Datapalooza Data Challenge

Licensure & Certification

- Generative AI with Large Language Model (DeepLearning.AI)
- Microsoft Certified Azure AI Fundamentals, and Data Fundamentals.
- Machine Learning (Coursera)
- Deep Learning Specialization (Coursera)
- Certified Base Programmer for SAS®9(License BP000759v9)
- Certified Advanced Programmer for SAS®9(License AP000202v9)
- Certified in Public Health for The National Board of Public Health Examiners (NBPHE)(License 6622)
- SAS Certified Clinical Trials Programmer Using SAS®9(License CTP000539v9)

PUBLICATIONS

1. Xin Hou, Yong He, Pan Fang, Shun Zhang, etc Using artificial intelligence to document the hidden RNA virosphere, Cell, Oct,2024, [https://www.cell.com/cell/fulltext/S0092-8674\(24\)01085-7](https://www.cell.com/cell/fulltext/S0092-8674(24)01085-7)
2. Yong He, Cheng Wang, Shun Zhang, Nan Li, Zhaorong Li, Zhenyu Zeng, KG-MTT-BERT: Knowledge Graph Enhanced BERT for Multi-Type Medical Text Classification, arXiv preprint arXiv:2210.03970
3. Yixi Xu, Raquel R Barbieri, Lucy Setian, Shun Zhang, et al., Reimagining leprosy elimination with AI analysis of a combination of skin lesion images with demographic and clinical data, The Lancet Regional Health-Americas,2022/5,(9)
4. Shun Zhang, George Rust, Kathryn Cardarelli, Jesus Felizzola, Mesfin Fransua, Harold G. Stringer, Jr.; Adherence to highly active antiretroviral therapy impact on clinical and economic outcomes for Medicaid enrollees with HIV and hepatitis C co-infection; AIDS Care, 2015, 27 (7), 829-835
5. Shun Zhang, Shanell L. McGoy, Daniel Dawes, Mesfin Fransua, George Rust, David Satcher, "The Potential for Elimination of Racial-Ethnic Disparities in HIV Treatment Initiation in the Medicaid Population among 14 Southern States", *PLOS One*; Apr 25th, 2014.
6. Shun Zhang, Charles Senteio, Jesus Felizzola, Rust George, Disparities in Antiretroviral treatment among Medicaid HIV-infected pregnant women, 2005-2007, *American Journal of Public Health*, December 2013, Vol. 103, No. 12, pp. e46-e53.
7. Shun Zhang, Kathryn Cardarelli, Ruth Shim, Jiali Ye, Karla L. Booker, Rust George, "Racial disparities in economic and clinical outcomes of pregnancy among Medicaid recipients", *Maternal and Child Health Journal* 2012;15:45-54.

8. Dominic Mack, Shun Zhang, Megan Douglas, Charles Sow, Harry Strothers, George Rust, Disparities in primary care Electronic Health Record (EHR) adoption rates, *Journal of health care for the poor and underserved* 27 (1), 327.
9. RS Shim, MT Compton, S Zhang, K Roberts, G Rust, BG Druss. Predictors of Mental Health Treatment Seeking and Engagement in a Community Mental Health Center, *Community Mental Health Journal* 2017, 53 (5), 510-514
10. George Rust, Shun Zhang, Zhongyuan Yu, Lee Caplan, Sanjay Jain, Turgay Ayer, Luceta McRoy, Robert Levine, Counties eliminating racial disparities in colorectal cancer mortality, *Cancer*, 2016 March.
11. Dominic Mack, Shun Zhang, Megan Douglas, Charles Sow, Harry Strothers, George Rust, Disparities in Primary Care EHR Adoption Rates. *Journal of Health Care for the Poor and Underserved*, Volume 27, Number 1, February 2016, pp. 327-338
12. Wonsuk Yoo, Geroge Rust, Shun Zhang, James Lillard, Health disparity characteristics on growth patterns of breast cancer mortality trends among the US Counties, 1989-2010, *Cancer Research* 75 (15 Supplement), 3701-3701
13. George Rust, Shun Zhang, Khusdeep Malhotra, Leroy Reese, Luceta McRoy, Peter Baltrus, Lee Caplan, Robert S Levine; Paths to Health Equity – Local Area Variation in Progress Toward Eliminating Breast Cancer Mortality Disparities, 1990-2009. *Cancer*, 2015, August, Volume 121, Issue 16, pages 2765–2774
14. George Rust, Shun Zhang, Luceta McRoy, Maria Pisu; Potential Savings from Increasing Adherence to Inhaled Corticosteroid Therapy in Medicaid-enrolled Children; *The American Journal of Managed Care*, 2015, March 21 (3), 173-180;
15. Abara, Winston E., Lerissa Smith, Shun Zhang, Amanda J. Fairchild, Harry J. Heiman, and George Rust. "The Influence of Race and Comorbidity on the Timely Initiation of Antiretroviral Therapy Among Older Persons Living With HIV/AIDS." *American Journal of Public Health*, (2014): e1-e7.
16. George Rust, Shun Zhang, Kelvin Holloway, Yasmin Tyler-Hill; "Timing of Emergency Department Visits for Childhood Asthma after Initial Inhaled Corticosteroid Use", *Population Health Management*, July 2014.
17. Khusdeep Malhotra, Peter Baltrus, Shun Zhang, Luceta McRoy, Lilly Cheng Immergluck, George Rust, "Geographic and Racial Variation in Asthma Prevalence and Emergency Department Use among Medicaid-Enrolled Children, in Fourteen Southern States", *Journal of Asthma*, June 2014, No. 10, pp 1-24.
18. Ruth S. Shim, Benjamin G. Druss, Shun Zhang, Giyeon Kim, Adesoji Oderinde, Sosunmolu Shoyinka, George Rust. Emergency Department Utilization among Medicaid Beneficiaries with Schizophrenia and Diabetes: The Consequences of Increasing Medical Complexity. *Schizophrenia Research*, Accept: November 20, 2013.
19. George Rust, Shun Zhang, Reynolds Joshua. Inhaled Corticosteroid Adherence and Emergency Department Utilization Among Medicaid-enrolled Children with Asthma, *Journal of Asthma*, September 2013, Vol. 50, No. 7, pp 769-775.
20. Fang Fang Zhang, Roberto Cardarelli, Joan Carroll, Shun Zhang, Kimberly G. Fulda, Karina Gonzalez, Jamboor K. Vishwanatha, Alfredo Morabia, Regina M. Santella; "Physical activity and global genomic DNA methylation in a cancer-free population", *Epigenetics* 6:3, pp 293-9, March 2011.
21. Charles Senteio, Summer Wright Collins, Rachael Jackson, Stacy Welk, Shun Zhang. "Effective Resources Supporting Healthy Sexual Behavior in Formerly Incarcerated Persons", *American Journal of Sexuality Education*, 2010, 5: pp 362–376.
22. Zhang Shun, Ding Xiaocang, Wang Haimin. The survey and analysis of impact on the financial crisis on the private medical services industry in Jing'an District, *Shanghai, Chinese Health Services Management*, No.7. July.2009, pp 475-476.
23. Zhang Shun, Ding Xiaocang, et al. "The dynamic monitoring survey of collected garbage spots in earthquake attacked area, Xuankou Township, Wenchuan County", *Chinese Primary Health Care*, Vol.22, No.275, Nov. 2008, pp 22-23.
24. Zhang Shun, Zhu Xiaozhen. "The survey on AIDS related knowledge among women who have work in the club of downtown Shanghai", *Chinese Journal of Public Health*, Vol.24, Suppl. Aug. 2008, pp 94-95.
25. Zhang Shun, Zhu Xiaozhen, et al. "Survey on attitude and behavior of sexual intercourse and knowledge of AIDS prevention among theatre academy students", *Chinese Journal of School Health*, Vol. 29, No.7-B, 2008, pp 15-16.
26. Zhu Xiaozhen, Zhang Shun, et al. "100 Questions on AIDS", ISBN 978-7-5072-1333-1.
27. Wang Liangfeng, Zhang Shun, et al, "Current Status of self-esteem of middle school students and influential factors in the rural area of Anhui Province", *Modern Preventive Medicine*, Vol.34, No.1 Jan. 2007, pp 25-27,35

28. Zhang Shun, Wang Liangfeng, et al, “Epidemiological survey on the current situation of social anxiety of left-behind children in primary school”, *Modern Preventive Medicine*, Vol.34, No.3 Feb. 2007, pp 441-443.
29. Zhang Shun, Wang Liangfeng, Sun Yehuan, et al. “Current status and its influential factors of children’s feelings of loneliness in the rural area of Anhui province”, *Chinese Journal of Disease Control and Prevention*, Vol.11, January 2007, pp 61-64.
30. Zhang Shun, Wang Liangfeng, Sun Yehuan et al “Life Events and Self-esteem of Middle School Students in Rural Area of Anhui Province”, *Chinese Mental Health Journal*, Vol.20, No.11, November 2006, pp 730-732.

Linkedin: <https://www.linkedin.com/in/shun-zhang-b3aaa026/>

Email: Lonver126@hotmail.com

Phone: 8178967968

Former Vice-President Joe Biden Awarded me the First Prize winner of Health Datapalooza in DC in 2016

