Shuyi Dong

+8613774193639 | sydong@smail.nju.edu.cn

EDUCATION

Nanjing University   Nanjing, China
M.S. in Information Science   *Aug. 2022 - June 2025 (expected)*

Wuhan University   Wuhan, China
B.S. in Information Management and Systems, GPA 3.84/4 (5/65)   *Aug. 2018 – June 2022*

PUBLICATIONS

- **Dong, S**., Mao, J., Ke, Q., & Pei, L. (In review). Decoding the Writing Styles of Disciplines: A Large-Scale Quantitative Analysis. Submitted to *Information Processing & Management* in August 2022.

- **Dong, S**., Mao, J., & Pei, L. (2023). Comparing the Writing Styles of Multiple Disciplines: A Large-Scale Quantitative Analysis. *Proceedings of the Association for Information Science and Technology*, 60(1), 941-943.

- Zhao, Y., Yang, B., **Dong, S**., et al. (2021). Design and Implementation of a Mobile Library Personalized Recommendation System Driven by Multi-Source Big Data. *Research on Library Science (Chinese Core)*, 2021(11), 20-31. https://doi.org/10.15941/j.cnki.issn1001-0424.2021.11.003

RESEARCH EXPERIENCE

Technology Trajectory Analysis Based On S&T Path-Linkage Perspective
Supervised by Prof.Zhichao Ba (NJU) and Prof.Lei Pei (NJU)   *May 2023 – Present*

- Discovered and evaluated how the interaction between science and technology contribute to the recombination in technology innovation
- Theorized five distinct interaction patterns between technologies and scientific outcomes, revealing their implications for technology innovation
- Use SemRep to extract medical concept
- Utilized Pajek to conduct the main path analysis letting the top global 15 key routes as roadmapping
- Utilized link prediction methods to effectively identify and forecast new combinations of technology components, achieving an AUC of 0.76
- Demonstrated the framework's efficacy through a case study in the anti-Leukemia field

Decoding the Writing Styles of Disciplines: A Large Scale Quantitative Analysis
Supervised by Prof.Jin Mao (WHU) and Prof.Qin Ke (CityU)   *June 2022 – Aug. 2023*

- Tracked evolution and differences of disciplinary writing styles over 30 years through quantitative analysis of 14M academic abstracts
- Constructed a feature framework of 14 features to measure the academic writing style
- Conducted a temporal t-test, pairwise t-test, and KW test to examine the differences in linguistic features across disciplines
- Employed machine learning to verify the discernibility of writing styles across various disciplines, achieving a robust classification performance with an AUC of 0.82
- Employed SHAP to identify key linguistic features, revealing linguistic disparities between soft and hard sciences

Construction of An Integrated Academic Lecture Data Platform
Supervised by Prof.Gang Li (WHU) and Prof.Jin Mao (WHU)   *March 2020 – Sept. 2021*

- Manually collected seed URLs of academic lecture webpages from various faculties of China's National Key Universities
- Built an academic report data platform for analyzing informal scientific communication, and exploring communication patterns
- Coded Python crawler to crawl academic reports of double first-class universities; configure Brat annotation platform for manual annotation
- Constructed academic communication relationship networks and exploring communication patterns

Mask Recognition System
Supervised by Prof.Andrew Lim (NUS) in the NUS summer research program of AI *June 2020 – Sept. 2020*
- Collected labeled face pictures with and without masks
- Read video streams and converting them into individual frames
- Built an MTCNN (Multi-task Cascaded Convolutional Networks) network to process the images. This involved scaling the images, determining the facial region, and identifying the positions of facial features

Biographical Text Mining and Application
Course work *May 2020 – July 2020*
- In-depth exploration of biographical text corpora and structured representation of text
- Utilized the LTP Named Entity Recognition module to label individuals in the text; conducted event extraction based on semantic role labeling
- Used the Word2Vec word vector model to represent character roles and built a character similarity calculation model
- Analyzed social networks of characters using Gephi; constructed a biographical knowledge graph based on Neo4j

WORK EXPERIENCE

eBay
Product Manager Intern *June 2022 – Sept. 2022 (Remote)*
- Streamlined the Onboarding Process for eBay Sellers
- Conducted research on onboarding processes for 14 of competitor companies, created comprehensive onboarding process diagrams using Figma, and generated reports outlining competitive advantages and disadvantages
- Modified the eBay onboarding process, specifically in the login phase, and reorganized the sequence of account creation phases.
- Revamped the welcome page and error state, and streamlined the mandatory information required from sellers for the risk model
- Created wireframes for the new process flowchart

SCHOLARSHIP

- **First-class Graduate Scholarship** of Nanjing University, 2023
- **First-class Scholarship** of Wuhan University, 2019 (Top 5% based on academic performance)
- **Second-class Scholarship** of Wuhan University, 2020 (Top 10% based on academic performance)
- **Third-class Scholarship** of Wuhan University, 2021 (Top 25% based on academic performance)
- **Merit student** of Wuhan University, 2019, 2020, 2021
- National Third Prize of 16th "Citi Cup" Financial Innovation Application Competition, 2021
- First Prize of Wuhan University Summer Social Practice Activity, 2020

TECHNICAL SKILLS

- **Programming**: Python, SQL, Java, STATA, SPSS, Matlab
- **Frameworks**: PyTorch, Scikit-Learn, NLTK, SciPy
- **Tools**: Git, Jupter Notebook, Visual Studio Code, Figma
- **Soft Skills**: Communication, Teamwork, Writing, Public Speaking, Critical Thinking
- **TOEFL**: 107 (Reading 28, Listening 30, Speaking 24, Writing 25)