Ke Yang

PhD Candidate, Computer Science, New York University

ky630@nyu.edu
+1 2159005854
Google Scholar page
github: KeYang0923
keyang0923.github.io

Education

Ph.D. in Computer Science, New York University   05.2021
- Advisor: Julia Stoyanovich, GPA 3.89
- Dissertation: Fairness, diversity, and interpretability of set-wise algorithmic outcomes
- Committee: Julia Stoyanovich (chair), Sebastian Schelter, Daniel Neill, and Oded Nov

Ph.D. in Computer Science, Drexel University   2015-2018
- Advisor: Julia Stoyanovich, GPA 3.98

M.S.E in Computer Science, Beijing Technology and Business University,   06.2015
- Advisors: Qian Mo and Zhongming Han, GPA 3.51
- Thesis: Detecting web spammers in online shopping websites
- Graduated summa cum laude

B.Eng. in Software Engineering, Beijing Technology and Business University   06.2012
- Advisor: Qian Mo, GPA 3.45
- Graduated summa cum laude

Research Interests

human-centered data science, data equity in Artificial Intelligence (AI) and Machine Learning (ML), and accountable data and model management systems.

- Develop efficient modeling techniques that account for data quality issues (e.g. biased representation in historical records and noises from heterogeneous sources), prompting users to be aware of the potential risks of the data-driven decisions.

- Evaluate and apply modeling techniques in AI and ML systems, helping users to interpret the systems' outputs without undue load to them.

- Build data and model management systems that highlight uncertainty and limitations of the data and model, equipping users to calibrate their trust in the systems.

Publications

Under submission

1. **Ke Yang**, Joshua R. Loftus, and Julia Stoyanovich, "Causal intersectionality for fair ranking"

Journals

1. Zhongming Han, **Ke Yang**, Xusheng Tan, "Analyzing Spectrum Features of Weight User Relation Graph to Identify Large Spammer Groups in Online Shopping Websites" in *Chinese Journal of Computers*, 40(4): 939-954 (2017) (in Chinese), DOI:10.11897/SP.J.1016.2017.00939

2. Zhongming Han, **Ke Yang**, Fengmin Xu, Dagao Duan, "Probabilistic Graphical Model for Detecting Spammers in Microblog Websites" in *Journal of Embedded System*, 8(1): 12-23 (2016), DOI:10.1504/IJES.2016.073747

3. Qian Mo, **Ke Yang**, "Overview of Web spammer detection" in *Journal of Software*, 25(7):

1505-1526 (2014) (in Chinese), DOI:10.13328/j.cnki.jos.004617

Peer-Reviewed Conference Proceedings

1. **Ke Yang**, Vasilis Gkatzelis, and Julia Stoyanovich, "Balanced ranking with diversity constraints", in *Proceedings of the 28th International Joint Conference on Artificial Intelligence* (IJCAI 2019), Macao, China, DOI:10.24963/ijcai.2019/836

2. Julia Stoyanovich, **Ke Yang**, HV Jagadish, "Online Set Selection with Fairness and Diversity Constraints" in *Proceedings of the 21th International Conference on Extending Database Technology* (EDBT 2018), Vienna, Austria, DOI:10.5441/002/edbt.2018.22

3. **Ke Yang**, Julia Stoyanovich, "Measuring Fairness in Ranked Outputs" in *Proceedings of the 29th International Conference on Scientific and Statistical Database Management* (SSDBM 2017), Chicago, USA, DOI:10.1145/3085504.3085526

Workshop Papers

1. **Ke Yang**, Biao Huang, Julia Stoyanovich, and Sebastian Schelter, "Fairness-Aware Instrumentation of Preprocessing Pipelines for Machine Learning", in *Proceedings of the 4th Workshop on Human-In-the-Loop Data Analytics, in conjunction with the 2020 ACM SIGMOD/PODS Conference* (HILDA@SIGMOD 2020), Portland, OR, USA

2. **Ke Yang**, Julia Stoyanovich, Abolfazl Asudeh, Bill Howe, HV Jagadish, and Gerome Miklau, "A Nutritional Label for Rankings" (demo) to appear in *Proceedings of the 2018 ACM International Conference on Management of Data* (SIGMOD 2018), Houston, USA

Patents

1. Kai Wang, **Ke Yang**, "A method and application system of detecting web spammers" in *China Invention Patent*, (2015) (in Chinese), No. CN201510012860

| Professional Experience | **Research Assistant**, Tandon School of Engineering, New York University | 2019-2021 |

Research Assistant, Tandon School of Engineering, New York University   2019-2021
Supervisor: Julia Stoyanovich
- Worked in Data, Responsibly to study the foundations of responsible data science and build tools that translate insights into data science practice. More details in dataresponsibly.github.io.

Research intern, AT&T Lab Research, New York   2019 Summer
Supervisors: Emily Dodwell, Ritwik Mitra, and Balachander Krishnamurthy
- Worked on a project to insure fairness and transparency for internal AT&T machine learning projects. This project aims to ensure that the projects' outcome meets reasonable guidelines for fairness without illegal bias and discrimination. We developed a diagnostic tool for bias-related issues that can be used by project managers and data scientists in AT&T.

Research Assistant, College of Computing & Informatics, Drexel University,   2015-2018
Supervisor: Julia Stoyanovich
- Worked in a project to model fair and diverse representation of individuals in subset selection and ranking process, and develop algorithms to ensure it with intersectional concerns in different applications.

Research engineer, Elite & Resource (start-up company)   2014-2015
Supervisor: Peng Sun
- Worked on a project to model historical data of flood disaster and develop techniques to help with flood prevention. In this project, we proposed a model to predict the probability of future

flood using water flow rate of small watershed torrents, which is recognized as a significant signal of potential flood. Our model is integrated as a core component of national watershed data management system.

Teaching & Advising Experience	**Lab section leader** at New York University - Special Topics in Data Science: Responsible Data Science, 2019 Spring **Teaching assistant** at Drexel University - Fundamentals of Databases (graduate level), 2017 Winter - Database Systems (undergraduate level), 2017 Spring **Research advisor** at New York University Independent Studies: - Jensine Raihan, 2020 Fall, "Ranking Facts in real datasets" - Biao Huang, 2020 Spring, "Fairness-aware instrumentation of preprocessing pipelines for machine learning" Research assistant: - Samasth Norway Ananda, 2020 Spring and Summer, "FairPrep: promoting data to a first-class citizen in studies on fairness enhancing interventions"
Academic Service	**Conference Organizing Committee:** - Technical chair: Workshop on Data Management for End-to-End Machine Learning (DEEM) at SIGMOD 2020 **Journal Review:** - Information Systems 2019 **Conference Proceedings Review:** - ACM Conference on Fairness, Accountability, and Transparency (ACM FAccT) 2020 - Workshop on Data Management for End-to-End Machine Learning (DEEM) at SIGMOD 2020 - ACM Conference on Human Factors in Computing Systems (CHI) 2020 **Student volunteer:** - ACM Conference on Fairness, Accountability, and Transparency (ACM FAT*) 2019
Talks	- Diversity in Set-wise Outcomes, 2019, AT&T Lab Research Weekly Symposium - Fairness In Ranked Outputs, 2018, AT&T Research Graduate Student Symposium
Fellowships, Awards, & Honors	- A selected participant at ACM FAT* PhD consortium, 2019 - A selected participant at AT&T graduate student symposium, 2018 - Invited participant at Fairness in ML Workshop at Google, 2018 - Student Scholarship at ACM Conference on Fairness, Accountability, and Transparency (ACM FAT*) 2018 - Fellowship for Doctoral Study, Graduate School, Beijing Technology and Business University,

2015

- Dissertation Award, Graduate School, Beijing Technology and Business University, 2015