Jinseok Kim

Research Assistant Professor
Institute for Research on Innovation & Science (IRIS)
Survey Research Center (SRC), Institute for Social Research (ISR)
School of Information, University of Michigan
330 Packard, Ann Arbor, 48104-2910
Phone: (734) 763-4994 | Email: jinseokk@umich.edu

Education

Ph.D. (2017): School of Information Sciences, University of Illinois at Urbana-Champaign
    Dissertation: *The Impact of Author Name Disambiguation on Knowledge Discovery from Large-Scale Scholarly Data*
    (Committee: Jana Diesner-Chair, Catherine Blake, Vetle Torvik, Michelle Shumate, and Seok-Hyoung Lee)
M.A. (2012): Department of Communication, University of Illinois at Urbana-Champaign
B.A. (2001): Department of English Language and Literature, Yonsei University, Korea

Appointment

2021.01 ~ present: Research Assistant Professor (by courtesy), School of Information, University of Michigan
2017.08 ~ present: Research Assistant Professor, IRIS, SRC, ISR, University of Michigan

Research Area

Machine Learning for Author Name Disambiguation, Data Linkage, Informetrics, and Collaboration Networks

Funded Research

PI OR CO-PI

National Science Foundation (Co-PI; 2019.9 ~ 2022.8; $462 K; PI-Jason Owen-Smith; 3.6 CM): *Creating a Data Quality Control Framework for Producing New Personnel-Based S&E Indicators.*

National Science Foundation (Co-PI; 2021.10 ~ 2022.9; $150 K; PI-Joelle Abramowitz; 2.0 CM): *FW-HTF-P Understanding Gig Work and its Effects on Wellbeing over the Life Course in the United States: A Machine Learning Approach*

Michigan Retirement and Disability Research Center (Co-PI; 2020.10 ~ 2021.9; $100 K; PI- Abramowitz, Joelle; 1.0 CM): *What We Talk about When We Talk About Self-Employment: Examining Self-Employment and the Transition to Retirement among Older Adults in the United States*

American Mathematical Society (PI: 2020.3~2020.11; $64 K; 2.0 CM): *Automating Author Name Disambiguation in MathSciNet*

Korea Institute of Science and Technology Information (Co-PI; 2020.4~2020.9; $27 K; PI-Jason Owen-Smith; 1.0 CM): *Analysis of AI-Based Service Technology for ScienceON*

INVESTIGATOR WITH SPECIFIED EFFORT

National Science Foundation (2021.5~2024.4; $1.28 M; PI- Jason Owen-Smith; 2.5 CM): *Linked NCSES Data for Evidence-Based Policy: Measuring Early Career Success and Modelling the Academic Research Supply Chain*

National Science Foundation (2020.5~2025.4; $990 K; UM PI- Jason Owen-Smith; 1.2 CM): *Collaborative Research: Impacts of Hard/Soft Skills on STEM Workforce Trajectories*

Publications (Jinseok Kim's name in bold; * = corresponding author)

JOURNAL ARTICLES

J21**: Kim, J.*,** Kim, J., & Kim, J. (May 2021, published online first). Effect of Chinese Characters on Machine Learning for Chinese Author Name Disambiguation: A Counterfactual Evaluation. *Journal of Information Science*.

J20: **Kim, J.*,** Kim, J., & Owen-Smith, J. (2021). Ethnicity-based Name Partitioning for Author Name Disambiguation Using Supervised Machine Learning. *Journal of the Association for Information Science and Technology. 72*(8), 979-994.

J19: **Kim, J.*,** & Owen-Smith, J. (2021). ORCID-Linked Labeled Data for Evaluating Author Name Disambgiuation at Scale*. Scientometrics. 126*(3), 2057-2083.

J18: **Kim, J.*,** & Owen-Smith, J. (2020). Model Reuse in Machine Learning for Author Name Disambiguation: An Exploration of Transfer Learning. *IEEE ACCESS. 8*, 188378-188389.

J17: **Kim, J.*,** & Kim, J. (2020). Effect of Forename String on Author Name Disambiguation. *Journal of the Association for Information Science and Technology. 71*(7), 839-855.

J16: **Kim, J.*** (2019). A Fast and Integrative Algorithm for Clustering Performance Evaluation in Author Name Disambiguation. *Scientometrics, 120*(2), 661-681.

J15: **Kim, J.*,** & Diesner, J. (2019). Formational Bounds of Link Prediction in Collaboration Networks. *Scientometrics, 119*(2), 687-706.

J14: **Kim, J.*** (2019). Scale-free Collaboration Networks: An Author Name Disambiguation Perspective. *Journal of the Association for Information Science and Technology, 70*(7), 685-700.

J13: **Kim, J.***, Kim, J., & Owen-Smith, J. (2019). Generating Automatically Labeled Data for Author Name Disambiguation: An Iterative Clustering Method. *Scientometrics, 118*(1), 253-280.

J12: **Kim, J.*** (2019). Author-Based Analysis of Conference Versus Journal Publication in Computer Science. *Journal of the Association for Information Science and Technology, 70*(1), 71-82.

J11: **Kim, J.*,** & Kim, J. (2018). The Impact of Imbalanced Training Data on Machine Learning for Author Name Disambiguation. *Scientometrics, 117*(1), 511-526.

J10: **Kim, J.*** (2018). Evaluating Author Name Disambiguation for Digital Libraries: A Case of DBLP. *Scientometrics, 116*(3), 1867-1886.

J09: **Kim, J.***, & Diesner, J. (2017). Over-Time Measurement of Triadic Closure in Coauthorship Networks. *Social Network Analysis and Mining, 7*(9)*.* 1-12.

J08: **Kim, J.***, & Diesner, J. (2016). Distortive Effects of Initial-Based Disambiguation on Measurements of Large-Scale Coauthorship Networks. *Journal of the Association for Information Science and Technology, 67*(6), 1446-1461.

J07: **Kim, J.***, Tao, L., Lee, S-H., & Diesner, J. (2016). Evolution and Structure of Scientific Co-Publishing Network in Korea between 1948 and 2011. *Scientometrics, 107*(1), 27-41.

J06: **Kim, J**.*, & Diesner, J. (2015). The Effect of Data Pre-processing on Understanding the Evolution of Collaboration Networks. *Journal of Informetrics, 9*(1), 226-236.

J05: **Kim, J.***, & Diesner, J. (2015). Coauthorship Networks: A Directed Network Approach Considering the Order and Number of Coauthors. *Journal of the Association for Information Science and Technology, 66*(12), 2685-2696.

J04: **Kim, J.***, & Kim, J. (2015). Rethinking the Comparison of Coauthorship Credit Allocation Schemes. *Journal of Informetrics, 9*(3), 667-673.

J03: **Kim, J**.*, & Diesner, J. (2014). A Network-Based Approach to Coauthorship Credit Allocation. *Scientometrics, 101*(1), 587-602.

J02: Diesner, J.*, **Kim, J**., & Pak, S. (2014). Computational Impact Assessment of Social Justice Documentaries. *Journal of Electronic Publishing, 17*(3). DOI: 10.3998/3336451.0017.306

J01: **Kim, J**., Kim, H., & Diesner, J.* (2014). The Impact of Name Ambiguity on Properties of Coauthorship Networks. *Journal of Information Science Theory and Practice, 2*(2), 6-15.

PEER-REVIEWED CONFERENCE PAPERS

C03: Diesner, J.*, Evans, C., and **Kim, J.** (2015) Impact of Entity Disambiguation Errors on Social Network Properties. International AAAI (Association for the Advancement of Artificial Intelligence) Conference on Web and Social Media (ICWSM). Oxford, UK.

C02**: Kim, J.***, Diesner, J., Aleyasen, A., Kim, H., & Kim, H.-M. (2014). Why Name Ambiguity Resolution Matters for Scholarly Big Data Research. 2014 IEEE Conference on Big Data, Washington, DC.

C01: Diesner, J.*, Pak, S., **Kim, J.**, Aleyasen, A., & Soltani, K. (2014). Computational Assessment of the Impact of Social Justice Documentaries. *iConference 2014*. Berlin, Germany.

PEER-REVIEWED EXTENDED ABSTRACTS & PRESENTATIONS AT PROFESSIONAL MEETINGS

P09: **Kim, J.,** & Diesner, J. (2018). Biases in Bibliometric Network Data and the Measurement of Triadic Closure. Presentation at the 2nd European Symposium on Societal Challenges in Computational Social Science: Discrimination and Bias, Cologne, Germany

P08: **Kim, J**., & Diesner, J. (2014). The Impact of Data Quality on Knowledge Discovery in Big Data: A Case of Name Ambiguity. Presentation at the Daegu-Gyeongbuk International Social Network Conference. Daegu, Korea.

P07: Diesner, J., & **Kim, J**. (2014). Computational Impact Assessment of Social Justice Media. Presentation at the Daegu-Gyeongbuk International Social Network Conference. Daegu, Korea.

P06**:** Diesner, J., **Kim, J**., & Higgins, A. (2014). Socio-Semantic Network Analysis for Impact Assessment. European Social Networks Conference (EUSN 2014). Barcelona, Spain.

P05**: Kim, J**., & Diesner, J. (2014). Structure of Scientific Collaboration: Considering the Order of Co-authors for Co-authorship Networks. Presentation at the International Network for Social Network Analysis. Saint Pete, FL.

P04**:** Diesner J, **Kim J**, Pak S (2014). Socio-Semantic Network Analysis for Impact Assessment of Documentaries. Presentations at International Network for Social Network Analysis. St Pete Beach, FL.

P03**:** Diesner, J., Aleyasen, A., **Kim, J.**, Mishra, S., & Soltani, S. (2013). Using Socio-Semantic Network Analysis for Assessing the Impact of Documentaries. WIN (Workshop on Information in Networks) 2013. New York, NY.

P02**:** Diesner J, Aleyasen A, **Kim J**, Mishra S, Soltani K, Wilner S (2013). Impact Assessment and ConText. Dissection: Impact. Oakland, CA.

P01: **Kim, J**., & Diesner, J. (2013). Structural Balance in Story Structures. Presentation at the International Network for Social Network Analysis. Hamburg, Germany.

Teaching

Instructor on data quality control and machine learning: (1) Total Data Quality (A Coursera certification program of 3 courses over 12 weeks; to be launched in May 2022): Co-instructor; developed two python packages to instruct data cleaning and machine-learning-based imputation (2) Responsive Survey Design (RSD) Program Summer Course (8 hrs.; June 2020, June 2021): Co-instructor of Total Data Quality (3) Applications of Statistical Modeling (4 hrs.; Dec 2020): guest instructor on machine learning application and algorithmic biases in deep learning

Graduate instructor (2017 Summer & 2016 Fall semesters): Social Media Analytics (LIS590SML; real-time online course) – newly developed by Jinseok Kim to instruct graduate students in the use of text and network mining techniques for testing social science theories on Facebook, Twitter, and Yelp data.

Teaching assistant: 2016 Spring: Social Computing (LIS590SF)

Guest Instructor on collecting, analyzing, and visualizing network and text data: (1) U. of Illinois courses: Digital Humanities (LIS590DH; 2013, 2014), Information Networks (LIS590IN; 2013), Social Computing (LIS590SF; 2012), and Network Analysis (LIS590NA; 2012~2015) (2) Conference workshops: Daegu-Gyeongbuk Int'l Social Network Conference (Daegu, Korea; 2014) and Int'l Network for Social Network Analysis (Saint Pete, FL; 2014)

Invited Talks

"The Impact of Author Name Disambiguation on Knowledge Discovery from Large-Scale Scholarly Data: A Case of Power-Law Distribution" (2017). Department of Interaction Science, Sungkyunkwan University, Seoul, Korea.

"Mining Social Media Data with NodeXL" (2017). Brian Lamb School of Communication, Purdue University, West Lafayette, IN.

"Hands-on Network Analysis" (2016). Computational Social Science Workshop, The Institute for Computing in Humanities, Arts, and Social Sciences (I-CHASS). University of Illinois at Urbana-Champaign, IL.

"Coauthor Order: A Network Approach" (2013). Korea Institute of Science & Technology Information (KISTI). Daejeon, Korea

"Seeing the Unseen: Detecting Social Ties from Archival Data" (2012). Illinois Institute of Technology, Chicago, IL.

Grant Proposal Reviewer: NSF (CISE IIS, 2019); Alfred P. Sloan Foundation

Ad-hoc Journal Reviewer (2014 ~ present): International Journal of Information Management; Journal of the Association for Information Science & Technology; Journal of Informetrics; Network Science; Physica A; PlosOne; Quantitative Science Studies; Scientometrics; Social Network Analysis & Mining

Awards and Honors

Berner-Nash Memorial Award for Outstanding Doctoral Dissertation (2018); The Elssesser Fellowship (2017); Microsoft Azure for Research Award (2016/06~2017/05); Eugene Garfield Doctoral Dissertation Fellowship by Beta Phi Mu[1] (2016~2017); Best Student Paper 2014 (1st Place Winner): Awarded at the annual meeting of ASIS&T 2014 (Seattle, WA) by the Association for Information Science and Technology SIG/Metrics

---

[1] The International Library and Information Studies Honor Society