

RENZHE YU

<https://renzheyu.com>

renzheyu@tc.columbia.edu ◊ +1 (212) 678-3429

460 Grace Dodge Hall, 525 West 120th Street, New York, NY 10027

EMPLOYMENT

Assistant Professor, Learning Analytics / Educational Data Mining 2022–present
Teachers College, Columbia University

AFFILIATIONS

Teachers College / Columbia University
Faculty Affiliate, Edmund W. Gordon Institute for Advanced Study 2025–present
Research Affiliate, Community College Research Center 2022–present
Faculty Member, Data Science Institute 2022–present

External

Research Affiliate, Pathways Network, Stanford University 2023–present
Research Affiliate, School of Education, University of California, Irvine 2022–present

EDUCATION

Ph.D. (Educational Data Science), University of California, Irvine 2022
M.A. (Economics of Education), Peking University 2017
B.S. (Artificial Intelligence), B.A. (Economics), Peking University 2014

PREVIOUS POSITIONS

Graduate Student Researcher, UC Irvine 2018–22
Research Intern, IBM Research 2020
Summer Fellow, Alan Turing Institute 2019
Visiting Researcher, UC Berkeley 2018
Graduate Student Researcher, Peking University 2014–17

AWARDS, HONORS & FELLOWSHIPS

Paper Awards

Best Full Paper Award, International Conference on Learning Analytics & Knowledge 2025
Best Full Paper Honorable Mention, International Conference on Educational Data Mining 2023
Best Undergraduate Paper, ACM Conference on Learning at Scale 2022
Best Paper Nomination, ACM Conference on Learning at Scale 2021
Best Paper Honorable Mention, AERA Educational Data Science Conference 2020
Outstanding Research Paper, Ministry of Education of China 2020
Best Paper Award, International Conference on Educational Data Mining 2018
Outstanding Research Paper, National Forum on Empirical Education Research 2018

Research Recognition

Emerging Scholar Award, Society for Learning Analytics Research	2023
National Academy of Education/Spencer Dissertation Fellowship (Declined)	2022
David P. Gardner Fellow, UC Berkeley	2021–22
Public Impact Fellow, UC Irvine	2021–22
Science for Social Good Fellow, IBM Research	2020
Data Science for Social Good Fellow, Alan Turing Institute	2019

Selective Events

Summer Institute in Computational Social Science, Princeton University	2021
Doctoral Colloquium, ACM Conference on Fairness, Accountability, and Transparency	2021,22
Early Career Cross-Community Young Researcher’s Event, London Festival of Learning	2018
LearnLab Summer School, Carnegie Mellon University	2018

Misc

Graduate Dean’s Fellowship, UC Irvine	2017
Outstanding Graduate, Peking University	2017
Leo KoGuan Scholarship, Peking University	2016
Merit Student, Peking University	2016
Travel Grant for Graduate Students, Peking University	2016
Academic Scholarship for Graduate Students, Peking University	2014,15
May 4th Scholarship, Peking University	2013,15

PEER-REVIEWED PUBLICATIONS

student mentee

† equal contribution or alphabetical order

* corresponding author

% acceptance rate, where known

Journal and Proceedings Articles

In computing-related fields, proceedings are primary venues for publication, equivalent to journals in other fields, with stringent, highly competitive peer-review process.

30. †Bird, K., †Castleman, B., †Song, Y., & †Yu, R. (In press). Is big data better? LMS data and predictive analytic performance in postsecondary education. *Journal of Research on Educational Effectiveness*.

29. Yao, C., Cortez, C., & Yu, R. (2025). Towards fair and privacy-aware transfer learning for educational predictive modeling: A case study on retention prediction in community colleges. In *Proceedings of the 15th International Conference on Learning Analytics & Knowledge (LAK '25)* (pp. 738–749). [30%]

• Best Full Paper Award

28. von Keyserlingk, L., Lauermann, F., Li, Q., Yu, R., Rubach, C., Arum, R., & Heckhausen, J. (2025). Students study activities before and after exam deadlines as predictors of performance in STEM courses: A multi-source data analysis. *Learning and Individual Differences*, 117, 102598.

27. Sabet, A., Bana, S., Yu, R., & Frank, M. (2024). Course-Skill Atlas: A national longitudinal dataset of skills taught in U.S. higher education curricula. *Nature Scientific Data*, 11(1086).

26. Lee, J., Hicke, Y., Yu, R., Brooks, C., & Kizilcec, R. (2024). The life cycle of large language models in education: A framework for understanding sources of bias. *British Journal of Educational Technology*, 55(5), 1982–2002.
25. †Xu, Z., †Olson, J., Pochinki, N., Zheng, Z., & Yu, R. (2024). Contexts matter but how? Course-level correlates of performance and fairness shift in predictive model transfer. In *Proceedings of the 14th International Conference on Learning Analytics & Knowledge (LAK '24)* (pp. 713–724). [30%]
24. Glandorf, D., Lee, H., Orona, G., Pumptow, M., †Yu, R., & †Fischer, C. (2024). Temporal and between-group variability in college dropout prediction. In *Proceedings of the 14th International Conference on Learning Analytics & Knowledge (LAK '24)* (pp. 486–497). [30%]
23. Chopra, H., Lin, Y., Samadi, M., Cavazos, J., Yu, R., Jaquay, S., & Nixon, N. (2023). Semantic topic chains for modeling temporality of themes in online student discussion forums. In *Proceedings of the 16th International Conference on Educational Data Mining (EDM '23)* (pp. 67–78). [26.5%]

- **Best Full Paper Honorable Mention**

22. Gardner, J., Yu, R., Nguyen, Q., Brooks, C., & Kizilcec, R. (2023). Cross-institutional transfer learning for educational models: Implications for model performance, fairness, and equity. In *Proceedings of the 6th Annual ACM Conference on Fairness, Accountability, and Transparency (FAccT '23)* (pp. 1664–1684). [26%]
21. Fischer, C., McPartlan, P., Orona, G., Yu, R., Xu, D., & Warschauer, M. (2022). Salient syllabi: Examining design characteristics of science online courses in higher education. *PLOS One*, 17(11), e0276839.
20. Moeller, J., von Keyserlingk, L., Spengler, M., Gaspard, H., Lee, H., Yamaguchi-Pedroza, K., Yu, R., Fischer, C., & Arum, R. (2022). Risk and protective factors of college students psychological well-being during the COVID-19 pandemic: Emotional stability, mental health, and household resources. *AERA Open*, 8.
19. Sabnis, S., Yu, R., & Kizilcec, R. (2022). Large-scale student data reveal sociodemographic gaps in procrastination behavior. In *Proceedings of the 9th ACM Conference on Learning at Scale (L@S '22)* (pp. 133–141). [30%]

- **Best Undergraduate Paper**

18. Umarji, O., Day, S., Xu, Y., Zargar, E., Yu, R., & Connor, C. (2021). Opening the black box: User-log analyses of childrens e-Book reading and associations with word knowledge. *Reading and Writing*, 34(3): 627–657.
17. Yu, R., Lee, H., & Kizilcec, R. (2021). Should college dropout prediction models include protected attributes? In *Proceedings of the 8th ACM Conference on Learning at Scale (L@S '21)* (pp. 91–100). [30%]

- **Best Paper Nomination**

16. Baker, R., Xu, D., Park, J., Yu, R., Li, Q., Cung, B., Fischer, C., Rodriguez, F., Warschauer, M., & Smyth, P. (2020). The benefits and caveats of using clickstream data to understand student self-regulatory behaviors: Opening the black box of learning processes. *International Journal of Educational Technology in Higher Education*, 17(13).

15. Fischer, C., Pardos, Z., Baker, R. S., Williams., J. J., Smyth, P., **Yu, R.**, Slater, S., Baker, R., & Warschauer, M. (2020). Mining big data in education: Affordances and challenges. *Review of Research in Education*, 44(1), 130–160.
14. **Yu, R.**, Li, Q., Fischer, C., Doroudi, S., & Xu, D. (2020). Towards accurate and fair prediction of college success: Evaluating different sources of student data. In *Proceedings of the 13th International Conference on Educational Data Mining (EDM '20)* (pp. 292–301). [**18.5%**]
13. Lin, Y., **Yu, R.**, & Dowell, N. (2020). LIWCs the same, not the same: Gendered linguistic signals of performance and experience in online STEM courses. In *Proceedings of the 21st International Conference on Artificial Intelligence in Education (AIED '20)* (pp. 333–345). [**26.6%**]
12. Rodriguez, F., **Yu, R.**, Park, J., Rivas, M., Warschauer, M., & Sato, B. (2019). Utilizing learning analytics to map students self-reported study strategies to click behaviors in STEM courses. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge (LAK '19)* (pp. 456–460). [**32%**]
11. Park, J., **Yu, R.**, Rodriguez, F., Baker, R., Smyth, P., & Warschauer, M. (2018). Understanding student procrastination via mixture models. In *Proceedings of the 11th International Conference on Educational Data Mining (EDM '18)* (pp. 187–197). [**16%**]

- **Best Paper Award**

10. Ha, W., & ***Yu, R.** (2017). How much is an improved school worth? Evidence from the comprehensive reform in compulsory education in Beijing. *Peking University Education Review*, 15(3), 137–153.
 - **Outstanding Research Paper** from the Ministry of Education of China
 - **Outstanding Research Paper** from the National Forum on Empirical Education Research
 - Translated version: Ha, W., & ***Yu, R.** (2019). Quasi-experimental evidence of a school equalization reform on housing prices in Beijing. *Chinese Education & Society*, 52(3-4), 162–185.
9. Ha, W., Wu, H., & **Yu, R.** (2015). A new research on the capitalization of school quality in housing prices: An empirical study based on repeated cross-sectional data in Beijing. *Education & Economy*, 05, 3–10.

Workshop and Work-in-Progress Papers

In computing-related fields, workshop and work-in-progress papers are published with stringent, selective peer-review process.

8. **Yu, R.**, Yang, H., Lin, X., Yao, C., Burkander, P., Thomas, K., & Mislavy, J. (2024). Technology-based instructional strategies show promise in improving self-regulated learning skills at broad-access postsecondary institutions. In *Proceedings of the 11th ACM Conference on Learning at Scale (L@S '24)* (pp. 408–411).
7. **Yu, R.**, Das, S., Gurajada, S., Varshney, K., Raghavan, H., & Lastra-Anadon, C. (2021). A research framework for understanding education-occupation alignment with NLP techniques. In *Proceedings of the 1st ACL Workshop on NLP for Positive Impact* (pp. 100–106).
6. Li, X., & **Yu, R.** (2021). Construction of weighted course co-enrollment network. In *Proceedings of the LAK Workshop on Using Network Science in Learning Analytics: Building Bridges towards a Common Agenda* (pp. 23–28).

5. Kung, C., & **Yu, R.** (2020). Interpretable models do not compromise accuracy or fairness in predicting college success. In *Proceedings of the 7th ACM Conference on Learning at Scale (L@S '20)* (pp. 413–416).
4. **Yu, R.**, Pardos, Z., & Scott, J. (2019). Student behavioral embeddings and their relationship to outcomes in a collaborative online course. In *Joint Proceedings of the Workshops of the 12th International Conference on Educational Data Mining (EDM '19)* (pp. 23–29).
3. **Yu, R.** (2019). Deconstructing the evolution of collaborative learning networks. In *Companion Proceedings of the 9th International Conference on Learning Analytics & Knowledge (LAK '19)* (pp. 741–745).
2. **Yu, R.**, Jiang, D., & Warschauer, M. (2018). Representing and predicting student navigational pathways in online college courses. In *Proceedings of the 5th ACM Conference on Learning at Scale (L@S '18)*.
 - **Spotlighted** at the Early Career Cross-Community Young Researcher's Event

Book Chapters

1. von Keyserlingk, L., Lauermann, F., **Yu, R.**, Rubach, C., & Arum, R. (2023). What can digital trace data tell us about postsecondary students academic success? An overview of the literature and an illustrative example. In *Jahrbuch der schulentwicklung. Band 23.* (pp. 128–149).

OTHER PUBLICATIONS

student mentee

† equal contribution or alphabetical order

Technical Reports

Acevedo, N., Cortez, C., Brooks, C., Kizilcec, R., & **Yu, R.** (2024). *Fairness Hub technical briefs: Mitigation strategies of distribution shift.* Learning Engineering Virtual Institute.

Acevedo, N., Cortez, C., Brooks, C., Kizilcec, R., & **Yu, R.** (2024). *Fairness Hub technical briefs: Definition and detection of distribution shift.* Learning Engineering Virtual Institute.

Lee, J., Brooks, C., **Yu, R.**, & Kizilcec, R. (2023). *Fairness Hub technical briefs: Overview of bias mitigation strategies.* Learning Engineering Virtual Institute.

Lee, J., Brooks, C., **Yu, R.**, & Kizilcec, R. (2023). *Fairness Hub technical briefs: AUC gap.* Learning Engineering Virtual Institute.

Lastra-Anadon, C., Das, S., Varshney, K., Raghavan, H., & **Yu, R.** (2021). *How universities can mind the skills gap. Higher education and the future of work.* Center for the Governance of Change, IE University.

Preprints / Working Papers

Yu, R., Xu, Z., Ch-Wang, S., & Arum, R. (2024). *Whose ChatGPT? Unveiling real-world educational inequalities introduced by large language models.* ArXiv.

Yu, R., Scott, J., & Pardos, Z. (2021). *Unsupervised representations predict popularity of peer-shared artifacts in online learning environment.* ArXiv.

- **Best Paper Honorable Mention** at the AERA Educational Data Science Conference

DATASETS AND SOFTWARE

† equal contribution or alphabetical order

†Calikus, E., †Trimarco, J., †Tseng, T., †**Yu, R.**, de Unanue, A., & Sipka, A. (2019). *Understanding and reducing inequities in transportation in the West Midlands* [Source code]. <https://github.com/alan-turing-institute/DSSG19-WMCA-PUBLIC>.

PRESENTATIONS

Invited Talks and Panels

AI in Education Research and Practice, Columbia AI Summit (Panelist)	2025
China Institute for Educational Finance Research, Peking University	2025
Department of Human Development, Teachers College, Columbia University	2024
AAAI Workshop on AI for Education	2024
Sociology of Algorithms Workshop, Columbia University	2024
Process Data Special Interest Group, Educational Testing Service	2024
Pathways Network, Stanford University	2023
China Institute for Educational Finance Research, Peking University	2023
Learning Analytics Research Network, New York University	2023
Waypoints Symposium on Data-Driven Student Success (Keynote Speaker)	2023
Center for Academic Innovation, University of Michigan	2023
Penn Center for Learning Analytics, University of Pennsylvania	2022
Department of Applied Psychology & Human Development, University of Toronto	2022
Department of Human Development, Teachers College, Columbia University	2022
Department of Information Systems, University of Maryland, Baltimore County	2022
School of Social Policy & Practice, University of Pennsylvania	2022
Department of Educational Psychology, Counseling, & Special Education, Penn State	2022
Department of Instructional Technology & Learning Sciences, Utah State University	2021
Colloquium on the Impact of AI on Ethics and Social Justice in Education (Panelist)	2021
IBM Research – Almaden Lab	2021
Digital Learning Lab, UC Irvine	2021
IBM Research	2020
eLearning Research and Practice Lab, Indiana University	2020
Future of Learning Lab, Cornell University	2019
Educational Technology Collective, University of Michigan	2019
School of Education, UC Irvine	2018

Non-Archival Conference Presentations (Selected)

Artificial Intelligence Research in Applied Linguistics Conference	2024
Columbia Data Science Day	2024
International Conference on Computational Social Science	2020–21,23–24
Annual Meeting of the American Educational Research Association	2019–20,22–24
Columbia Workshop on Fairness in Operations and AI	2023
Trustworthy AI Lab for Education Summit	2023
Unizin Summit	2022
Research on Academic Pathways Flash Conference	2021
AI for Social Good Workshop	2021
AERA Educational Data Science Conference	2020

Mechanism Design for Social Good Workshop	2020
Digital Learning in the Humanities and Beyond Symposium	2018
Annual Conference of the Association for Education Finance and Policy	2016–17
Annual Conference of the Association for Public Policy Analysis and Management	2016
Young Scholars’ Forum on Public Economics	2015
China’s Annual Conference on Economics of Education	2015

RESEARCH GRANTS

Principal Investigator , Teachers College Faculty Collaboration Grant <i>Understanding Postsecondary Student Success Through Curricular Content Analytics</i> Co-PI: Minaya Total award: \$20,000	2024–25
Co-Principal Investigator , Learning Engineering Virtual Institute <i>Fairness Analysis & Transfer Learning Hub</i> PI: Kizilcec; Co-PIs: Brooks, Yu Total award: \$475,000	2023–25
Co-Principal Investigator , Ascendium Education Group <i>Evaluations of Strategies to Improve Education and Workforce Success</i> PI: Castleman; Co-PIs: Bird, Finnegan, Yu Total award: \$1,276,000	2022–26
Principal Investigator , Office of Institutional Research, Peking University <i>Effects of Seminar Classes on College Students’ Creativity</i>	2014–15

TEACHING

Teachers College, Columbia University	
HUDK 5100: Supervised Research and Practice (Supervisor)	Fall 2022–present
HUDK 5053: Learning Analytics Practicum (Instructor)	Every Summer, 2023–present
HUDK 4051: Learning Analytics: Process and Theory (Instructor)	Every Spring, 2023–present
UC Irvine	
INF4MATX 199: Individual Study (Mentor)	Winter 2021–Spring 2022
EDUC 30: 21st Century Literacies (TA)	Spring 2019
EDUC 10: Educational Research Design (TA)	Winter 2019
R Workshop for Educational Data Science (Co-Instructor)	Winter 2019
Peking University	
State and Education (TA)	Spring 2016
Advanced Econometrics (TA)	Spring 2016
Guest Lectures	
COS 597H: AI, Society, and Education (Princeton University)	Fall 2024
B9153: Generative AI: Technical and Social (Columbia University)	Fall 2024
INF 1344H: Introduction to Statistics for Data Science (University of Toronto)	Fall 2022
EDUC 224B: Learning Analytics Practicum (UC Irvine)	Winter 2021
EDUC 30: 21st Century Literacies (UC Irvine)	Fall 2019

ACADEMIC SERVICE

Teachers College / Columbia University

Capstone Project Coordinator, M.S. in Learning Analytics Program	2023–present
Co-Organizer, PhD Working Group, Department of Human Development	2023–present
Capstone Project Sponsor, M.S. in Data Science Program	2024
Mentor, Data Science Institute Scholars Program	2024
Search Committee Member, Assistant/Associate Professor in Economics and Education (2)	2022–23

Professional Communities

Extended Steering Committee Member, ACM Conference on Learning at Scale	2023–present
Area Chair, ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization	2024
Publicity Co-Chair, ACM Conference on Fairness, Accountability, and Transparency	2023,24
Mentorship Program Lead, Mechanism Design for Social Good	2022–23
Co-Chair, Workshop on Fairness, Accountability, and Transparency in Educational Data	2020,22
Mentorship Co-Chair, ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization	2021
Co-Organizer, Digital Learning in the Humanities and Beyond Symposium	2018

Manuscript / Grant Proposal Review

ACM Conference on Learning at Scale	2018–19,25
ACM Conference on Fairness, Accountability, and Transparency	2024–25
Tools Competition	2025
The ACM Web Conference	2025
Nature Communications	2024
AAAI/ACM Conference on AI, Ethics, and Society	2024
British Journal of Educational Technology	2024
International Conference on Educational Data Mining	2019,24
International Learning Analytics & Knowledge Conference	2020–24
IEEE Internet Computing	2023
The Internet and Higher Education	2023
Journal of Educational Data Mining	2023
Journal of Experimental Psychology: Applied	2023
International Journal of Artificial Intelligence in Education	2022,23
ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization	2021–23
IEEE Transactions on Learning Technologies	2021,23
ACM Conference on Human Factors in Computing Systems	2019,21,23
AERA Open	2021–22
Journal of Learning Analytics	2020
International Conference of the Learning Sciences	2018

Student Roles

Student Volunteer, ACM Conference on Fairness, Accountability, and Transparency	2021–22
PhD Recruitment Student Buddy, School of Education, UC Irvine	2020
Organizer, UC Irvine Learning Analytics Reading Group	2019
Search Committee Member, Assistant/Associate Professor of Learning Analytics, UC Irvine	2018–19
Vice President, IEEE Peking University Student Branch	2012

MEDIA COVERAGE

Report: Teaching online students to self-regulate learning. (2024, August). *Inside Higher Ed*.

Here's how data can help unlock education equity. (2024, April). *Teachers College Newsroom*.

Shared modeling can help schools predict, avert dropouts. (2023, June). *Cornell Chronicle*.

How to make a campus a living laboratory to study learning. (2023, May). *The Chronicle of Higher Education*.

Navigating the risks and rewards of ChatGPT. (2023, January). *Teachers College Newsroom*.

Can analyzing clicks in digital systems predict which students are struggling? It depends. (2022, October). *EdSurge*.

Marginalized students suffer penalties from procrastination. (2022, June). *Cornell Chronicle*.

Testing AI fairness in predicting college dropout rate. (2021, June). *Cornell Chronicle*.

Questioning the role of AI in exam marking. (2021, April). *Raconteur*.

Good grades, stressed students. (2021, March). *The Chronicle of Higher Education*.

As the pandemic upends higher education, is residential college worth the cost? (2020, July). *PBS News Hour*.