Rohan Alexander

Assistant Professor, Information and Statistical Sciences, University of Toronto
Room 714, 140 St George St, Toronto, Ontario, M5S 3G6, Canada
✉ rohan.alexander@utoronto.ca ⦿ RohanAlexander 🐦 rohanalexander 🌐 rohanalexander.com

Academic appointments

- Assistant Professor, Information & Statistical Sciences, University of Toronto   July 2020 - *Current*
  - Parental leave: From June to August 2021 I was on parental leave.
  - COVID-19: From March to December 2020 I shared full-time childcare duties (0.5 FTE on average). From January 2021 to May 2021, and again from September to December 2021, I shared part-time childcare duties, (0.25 FTE on average).
- Assistant Director, Canadian Statistical Sciences Institute (CANSSI) Ontario September 2021 - *Current*
- Senior Fellow, Massey College   September 2021 - *Current*
- Academic Lead, Short courses, Data Sciences Institute   July 2021 - *Current*
  - The DSI offers eight OSAP eligible short courses, each of around 20-30 contact hours.
  - I led liaison with industry and academic partners; developed curricula; hired and managed a team of nine to create slides, assessment, and other teaching materials.
  - With staff support I have responsibility for hiring and managing instructors and TAs to teach the courses and improve the open-source materials.
- Faculty Affiliate, Schwartz Reisman Institute for Technology and Society   August 2020 - *Current*
- Co-lead, Data Sciences Institute Thematic Program in Reproducibility   November 2021 - June 2023
- Post-doctoral Fellow, Information, supervised by Kelly Lyons   July 2019 - July 2020

Education

- PhD (Economics), Australian National University   February 2015 - December 2019
  - Dissertation title: *Essays on Applied Historical Labour Economics*.
  - Committee: John Tang (chair); Martine Mariotti; Tim Hatton; and Zach Ward.
- M. Economics, Australian National University   February 2012 - November 2012
- B. Economics, University of Queensland   February 2004 - November 2008

Book

- Alexander, Rohan, 2023, *Telling Stories with Data*, Chapman and Hall/CRC.

Edited book

- Kennedy, Lauren, Rohan Alexander, and Andrew Gelman, (editors), *Multilevel Regression and Post-stratification: A Practical Guide and New Developments*, (accepted by Cambridge University Press in July 2021, expected publication in 2024).

Peer-reviewed journal articles[1]

- Katz, Lindsay, and Rohan Alexander, 2023, "Digitization of the Australian Parliamentary Debates, 1998-2022", *Scientific Data*, Accepted.
- Collins, Annie, and Rohan Alexander, 2022, "Reproducibility of COVID-19 pre-prints", *Scientometrics*, 4 July, https://doi.org/10.1007/s11192-022-04418-2.
- Alexander, Rohan, and Zach Ward, 2018, "Age At Arrival and Assimilation In The Age Of Mass Migration", *The Journal of Economic History*, 78(3), pp. 904-937, https://doi.org/10.1017/S0022050718000335.

---

[1]My work is interdisciplinary and there are different norms for author order and publication venue across disciplines. With regard to author order, Alexander and Ward, 2018, is alphabetical, all others follow the norm of the senior author last. Items in the "peer-reviewed conference proceedings" list—despite being a "conference proceeding"—can be "translated" as equivalent to a peer-reviewed journal article.

Peer-reviewed conference proceedings

- Ferguson, Sharon, Paula Akemi Aoyagui, Rohan Alexander, Anastasia Kuzminykh, 2022, "A Thematic Comparison of Human and AI Explanations of Sexism Assessment", *Human Centered AI Workshop at NeurIPS*.

Other publications

- Horton, Nicholas, Rohan Alexander, Aneta Piekut, and Colin Rundel, 2022, "Editorial: The Growing Importance of Reproducibility and Responsible Workflow in the Data Science and Statistics Curriculum", *Journal of Statistics and Data Science Education*, https://doi.org/10.1080/26939169.2022.2141001.

Pre-prints

- Kennedy, Lauren, Jonah Gabry, Dewi Amaliah, and Rohan Alexander, 2023, "mrpkit: A reproducible, opinionated, workflow for multilevel regression with post-stratification in R".
- Collins, Annie, and Rohan Alexander, 2022, "History of MRP", (under review at Kennedy, Alexander, and Gelman, (editors), *Multilevel Regression and Poststratification: A Practical Guide and New Developments*).
- Chiu, Ke-Li, Annie Collins, and Rohan Alexander, 2022, "Detecting Hate Speech with GPT-3", *arXiv*, 24 March, https://arxiv.org/abs/2103.12407.
- Alexander, Rohan, and Monica Alexander, 2021, "The Increased Effect of Elections and Changing Prime Ministers on Topics Discussed in the Australian Federal Parliament between 1901 and 2018", *arXiv*, 17 November, https://arxiv.org/abs/2111.09299.
- Alexander, Rohan, Samantha-Jo Caetano, Haoluan Chen, Michael Chong, Annie Collins, Shirley Deng, Isaac Ehrlich, Paul Hodgetts, Yena Joo, Marija Pejcinovska, Mariam Walaa, and Matthew Wankiewicz, 2021, "An Introduction to DoSStoolkit", *arXiv*, 19 May, https://arxiv.org/abs/2105.09347.
- Chiu, Ke-Li, and Rohan Alexander, 2021, "On consistency scores in text data with an implementation in R", *arXiv*, 13 January, https://arxiv.org/abs/2101.05225.
- Hodgetts, Paul A., and Rohan Alexander, 2020, "cesR: An R package for the Canadian Election Study", *SocArXiv*, 3 September, https://osf.io/preprints/socarxiv/a29h8/.

R packages

- Alexander, Rohan, and Paul A. Hodgetts, 2021, `AustralianPoliticians`.
- Hodgetts, Paul A., and Rohan Alexander, 2021, `cesR`.
- Alexander, Rohan, and A Mahfouz, 2021, `heapsofpapers`.

Awards

- 2022 ANU Australian Politics Studies Centre Visiting Scholar ($2,500)
- Honourable Mention, 2021 Master of Information Student Council Instructor Awards

Grants

- Mitacs Accelerate with the Investigative Journalism Foundation of Canada   July 2023 - June 2026
  - $426,666 (Mitacs: $235K, IJF: $192K) to support interns associated with "Follow the Money: Tracking the Influence of Lobbyists in Canada".
  - Supervisor: Rohan Alexander
  - Co-Supervisor: Christopher Cochrane
- Data Sciences Institute (DSI) Catalyst Grant   February 2023 - January 2024
  - $100,000 for "Providing data to improve representation by public officials".
  - Investigators: Peter Loewen (Named PI), Rohan Alexander (co-PI), Aya Mitani (co-PI), Elena Tuzhilina (co-PI).
- CANSSI Ontario Data Access Grant   January 2023
  - $5,000 to purchase data in support "Developing a large-scale lobbying and political donations dashboard to foster access to and interdisciplinary research on democratic accountability and political donations".

- – Co-PIs: Rohan Alexander and Chris Cochrane.
- Australian Research Council Discovery Projects   January 2023 - December 2025
  - – $287,146 for "Australian Parliamentary Speech: How Deliberative? How Representative?"
  - – Investigators: Marija Taflaga (Chief Investigator), Keith Dowding (Chief Investigator), Kenneth Benoit (Partner Investigator), Rohan Alexander (Partner Investigator).
- U of T Data Sciences Institute Data Access Grant   November 2022 - November 2023
  - – $9,993 to purchase data in support of the project "Developing a large-scale lobbying and political donations dashboard to foster access to and interdisciplinary research on democratic accountability and political donations".
  - – Co-PIs: Rohan Alexander and Chris Cochrane.
- SSHRC Connection Grant   October 2022 - September 2024
  - – $50,000 on the project "Follow the money: Mobilizing Canadian political donations data".
  - – PI: Rohan Alexander.
  - – Co-investigators: Jacob Robbins-Kanter, Wing Cheung Michael Wong, Zane Schwartz.
- SSHRC Partnership Development Grant   June 2022 - June 2024
  - – $169,855 on the project "Follow the Money: Donors, Donations, and Parties in Canadian Politics".
  - – PI: Christopher Cochrane.
  - – Co-applicant: Harold Jansen, Jacob Robbins-Kanter, Michael Pereira, Randy Besco, Rohan Alexander, Michael Wing-Cheung Wong, and Zane Schwartz.
- Mitacs Accelerate with the Investigative Journalism Foundation of Canada   June 2022 - November 2022
  - – $48,000 to support an intern associated with "Follow the Money: Investigating Political Donations in Canada".
  - – Co-PIs: Rohan Alexander and April Lindgren.
- U of T Data Sciences Institute Summer Undergraduate Data Science   May 2022 - August 2022
  - – $7,500 to support an undergraduate over summer on the project "Disparities in climate-induced health outcomes in the Greater Toronto Area".
  - – Co-PIs: Rohan Alexander and Monica Alexander.
- U of T School of Cities Urban Challenge Grant   January 2022 - December 2022
  - – $75,000 toward RAs and expenses associated with the project "Disparities in climate-induced health outcomes in the Greater Toronto Area".
  - – Co-PIs: Rohan Alexander, Monica Alexander, Karen Smith, and Micah Hewer.
- U of T Data Sciences Institute Data Access Grant   December 2021 - April 2022
  - – $10,000 to purchase ICES data in support of the project "Disparities in climate-induced health outcomes in the Greater Toronto Area".
  - – Co-PIs: Rohan Alexander and Monica Alexander.
- U of T Faculty of Information SSHRC Institutional Grant   November 2021 - March 2022
  - – $3,500 to support an RA on the project "At what cost? Using Hansard data to understand the public policy that went 'missing' due to COVID-19".
  - – PI: Rohan Alexander.
- U of T Faculty of Arts & Science Pedagogical Innovation and Experimentation Fund September 2021 - March 2022
  - – $10,000 to continue developing the DoSS toolkit.
  - – Co-PIs: Rohan Alexander, and Samantha-Jo Caetano.
- CANSSI Ontario Workshop/Lecture (COWL) Program   February 2021
  - – $1,000 to support undergraduate research.
  - – PI: Rohan Alexander.
- CANSSI Ontario Multidisciplinary Doctoral (Mdoc) training program   September 2021 - September 2025
  - – $10,000 annually for five years to support a PhD student.
  - – Co-PIs: Rohan Alexander, and Monica Alexander.
- Mitacs Accelerate with Sensibill   June 2021 - November 2021
  - – $30,000 to support an intern associated with "Robust Taxonomy out of Receipt Item Labels"
  - – Co-PIs: Rohan Alexander and Michael Guerzhoy.

- U of T Faculty of Arts & Science Pedagogical Innovation and Experimentation Fund September 2020 - March 2021.
  - $15,000 to develop an applied statistics toolkit.
  - Co-PIs: Rohan Alexander, and Samantha-Jo Caetano.
- UCL & U of T Strategic Partner Funds for AI collaborations   July 2020 - December 2020
  - $25,000 for "AI Governance: Building ethical frameworks to balance risk and innovation"
  - Co-PIs: Elizabeth Lomas, Kelly Lyons, Jenny Bunn, and Rohan Alexander.

Teaching experience

PHD

- Instructor, "INF3104: Data Science Foundations", Information   Winter 2023

MASTERS

- Instructor, "INF2178: Experimental Design for Data Science", Information   Winter 2021
- Instructor, "INF2178: Experimental Design for Data Science", Information   Winter 2020

UNDERGRADUATE

- *(Upcoming)* Instructor, "STA302: Methods of Data Analysis I", Statistical Sciences   Winter 2024
- *(Upcoming)* Instructor, "INF312: Worlds Become Data", Information   Winter 2024
- Instructor, "INF312: Worlds Become Data", Information   Winter 2023
- Instructor, "INF312: Worlds Become Data", Information   Winter 2022
- Instructor, "STA304: Surveys, Sampling and Observational Data", Statistical Sciences   Winter 2022
- Instructor, "STA304: Surveys, Sampling and Observational Data", Statistical Sciences   Fall 2020

Supervisory experience

PHD

- Ciara Zogheib, Information   September 2022 - *Current*

MASTERS

- Inessa De Angelis, Information (thesis supervisor)   September 2023 - *Current*
  - Thesis: *Torrential Twitter: Climate Change, Female Politicians, and Harassment*
- Lorena Almaraz De La Garza, Information (academic supervisor)   June 2022 - September 2022
  - Research paper: *Machine Learning Inference for Demographic Features of Political Donors*
  - First placement: Senior AI/ML Model Steward Partner, Thomson Reuters
- Ciara Zogheib, Information (thesis supervisor)   September 2021 - June 2022
  - Thesis: *The "Paper and Pencil" Operations of Interdisciplinarity: Interdisciplinary Integration as Information Practice*
  - First placement: PhD student, University of Toronto
- Elisa Du, Applied Computing (academic supervisor)   March 2021 - December 2021
  - Thesis: *Enriched Understanding of Retail Receipts for Personalized Financial Insights*
  - First placement: Machine Learning Engineer, Sensibill
- Thomas William Rosenthal, Information (academic supervisor)   January 2021 - April 2021
  - Research paper: *A Machine Learning Exploration into the Attributes that Define Single-Origin Coffee*
  - First placement: Data Scientist, Purpose Analytics

UNDERGRADUATE

- Connor Glossop, Engineering (senior thesis)   September 2022 - April 2023
  - Thesis: *Analyzing Political Polarization in Canada"s Parliament*
  - First placement: Data Scientist, Intact Insurance
- Callie Moore, Engineering (senior thesis)   September 2021 - April 2022
  - Thesis: *Dyadic Representation in Canadian Parliament*

– First placement: Data Analyst, Hospital for Sick Children
- Jeremy Hunt, Engineering (senior thesis)   September 2020 - April 2021
    – Thesis: *Study of Software Tool Usage in Academics*

Service

Professional societies

- President-elect, *Data Science and Analytics Section, Statistical Society of Canada*   July 2023 - June 2024

Editorial

- Associate Editor, *Journal of Statistics and Data Science Education*   Summer 2021 - *Current*

Refereeing

- 2023: *Journal of Statistics and Data Science Education* (x2); *Social Sciences and Humanities Research Council of Canada*; *Archival Science*; *Harvard Data Science Review*.
- 2022: *Chapman and Hall/CRC*; *Journal of Statistics and Data Science Education* (x3); *Political Analysis*; *Records Management Journal*.
- 2021: *Chapman and Hall/CRC* (x3); *Journal of Statistics and Data Science Education* (x2).
- 2020: *Canadian Journal of Political Science* (x2); *Political Analysis*; *Harvard Data Science Review*; *The Economic Record*.

Outreach

- Co-organizer, *Toronto Data Workshop*   September 2019 - *Current*
- Organizer, *1st Colloquium on Applications of Large Language Models*   11 April 2023
- Organizer, *3rd Toronto Conference on Reproducibility*   23 February 2023
- Local host, *Toronto Replication Games*   22 February 2023
- Organizer, *Florence Nightingale Day*   18 February 2023
- Organizer, *CANSSI Ontario Statistical Software Conference*   10 November 2022
- Organizer, *Florence Nightingale Day*   9 April 2022
- Organizer, *2nd Toronto Workshop on Reproducibility*   23-25 February 2022
- Organizer, *1st Toronto Workshop on Reproducibility*   25-26 February 2021
- Co-organizer, *MRP Working Group*   January 2020 - November 2021

Departmental and University Committees

- Member, Undergraduate Advisory Group, Information   July 2023 - *Current*
- Member, Data Sciences Speaker Series Selection Committee, Data Sciences Institute   March 2023 - *Current*
- Coordinator, Applied Research and Education Seminar Series, Statistical Sciences July 2022 - *Current*
- Member, Graduate Committee, Statistical Sciences   July 2020 - *Current*
- Member, BI Recruitment and Admission Committee, Information   July 2021 - June 2023
- Member, Advisory Committee, Dean Search, Faculty of Information   November 2022 - May 2023
- Co-organizer, *Statistical Sciences PhD Practice Job Talk Day*   12 December 2022
- Chair, Data Sciences Speaker Series Selection Committee, Data Sciences Institute March 2022 - March 2023
- Member, Search Committee, Dalla Lana School of Public Health, Assistant Professor—Data Science Methods for Population Health and Health Systems   July 2022 - March 2023
- Member, Faculty of Arts and Sciences Computation and Data Science Education Working Group September 2021 - July 2022
- Member, Applied Research and Education Seminar Series, Statistical Sciences July 2020 - June 2022
- Advisor, *Statistical Sciences PhD Student Research Day*   25 May 2022
- Co-organizer, *Statistical Sciences PhD Practice Job Talk Day*   17 December 2021
- Member, U of T Library Technical Skills Outreach Project   July 2021 - June 2022

- Member, Search Committee, Information, HCDS Teaching Stream Assistant Professor, November 2021 - April 2022
- Member, Search Committee, Information, Learning Hub Director and Teaching Stream Assistant Professor   November 2020 - April 2020
- Member, Awards Committee, Information   July 2020 - November 2020

Thesis Committees

- Member, Criminology, PhD Thesis Committee, Jamie Duncan   January 2023 - *Current*
  - Title: *The border as a platform*
- Member, Information, PhD Thesis Committee, Mirai Gendi   June 2022 - *Current*
  - Title: *Data science and process mining with application to digital forensics*

Exam Committees

- Examiner, Statistical Sciences, PhD Applied Research Comprehensive, Fred Song Summer/Fall 2023
- Member, Statistical Sciences, PhD Final Oral Examination, Blair Bilodeau   14 April 2023
- Chair, Statistical Sciences, PhD Oral Exam Committee, Michaël Lalancette   14 July 2022
- Second Reader, Information, MI Thesis Defense, June Lim   9 September 2021
  - Title: *Gendered Voices in Japanese Popular Music: A Data-Driven Analysis*
- Member, Statistical Sciences, PhD Oral Exam Committee, Alex Stringer   27 April 2021
- Member, Statistical Sciences, PhD Oral Exam Committee, Arvind Shrivats   12 February 2021
- Member, Statistical Sciences, PhD Oral Exam Committee, Zhen Qin   12 January 2021

Presentations

- *From teaching to practice: Insights from the Toronto Reproducibility Conferences*
  - 8 August 2023, Joint Statistical Meetings (JSM), Toronto.
- *Turning Our World Into Data*
  - 19 July 2023, Harvard Biostatistics Data Science in Action: Machine Learning for Self-driving cars (online).
  - 7 July 2021, Harvard Biostatistics Data Science in Action Summer Camp (online).
- *The Making of a Nation: Who Voted for Australian Federation?*
  - 23 June 2023, Australia before Federation, Centre for Economic History, Canberra.
- *Climate justice and credible data science*
  - 7 March 2023, U of T, Faculty of Information, Guest lecture, "Communities and Values".
- *Worlds become data*
  - 16 February 2023, U of T, Faculty of Medicine, Adventures In Science Mentorship Program.
- *Testing, testing, is this thing right?*
  - 27 January 2023, Quantitative Methods Cluster, University of Toronto.
  - 30 November 2022, FCT Abuja, Nigeria, UseR Group, (online).
  - 9 November 2022, Workshop in Computational & Quantitative Social Sciences, Toronto.
- *The growing importance of reproducibility and responsible workflow in the data science and statistics curriculum*
  - 13 December 2022, Panelist, CAUSE/JSDSE Webinar Series (online).
- *Reproducible workflows*
  - 2 November 2022, Biostatistics seminar series, Centre for Outcomes Research, McGill University Health Centre – Research Institute (online).
- *Rethinking data science*
  - 6 October 2022, Toronto Data Workshop, Toronto.
- *On the role of reproducibility in data science*
  - 4 August 2022, Stanford Data Science (online).
- *Integrating reproducibility into undergraduate data science teaching*
  - 14 July 2022, Australasian Teaching Economics Conference, (online).
  - 4 June 2022, Canadian Economics Association Conference, Carleton University, Ottawa.
- *On the double burden of disadvantage in data science*
  - 17 June 2022, We All Count, Talking Data Equity (online).

- *Reproducible Data Science Workflows in R*
  - 11 January 2022, Accra R User Ghana R Conference (online)
- *What data science can learn from the James Webb Space Telescope*
  - 17 December 2021, University of Toronto Statistics and MachIne LEarning (SMILE) Journal Club (online).
- *On the privilege of turning our world into data*
  - 16 December 2021, Young Irish Statisticians (online).
- *A super opinionated talk about data science, APIs, and grad school applications*
  - 18 November 2021, University of Toronto, Undergraduate Statistical Sciences Union Workshop (online).
- *Remarks on data science*
  - 19 October 2021, University of Toronto, Trinity College High Table (online).
- *Introduction to Andrew Gelman*
  - 17 September 2021, University of Toronto, Data Sciences Institute launch (online).
- *Introduction to R*
  - 22 December 2020, University of Toronto, Global Society for Genetics and Genome Biology (online).
- *Choosing the optimal set of co-variates to satisfy competing priorities in multi-level regression with post-stratification*
  - 6 May 2020, MRP Reading Group (online).
- *Forecasting Parliamentary Elections*
  - 10 December 2019, Monash University, Department of Econometrics & Business Statistics, Melbourne.
  - 9 December 2019, Australian Society for Quantitative Political Science Conference, Melbourne.
  - 2 December 2019, Australian National University, Cake for Comments, Canberra.
  - 20 November 2019, POP Politics Workshop, Canberra.
  - 17 October 2018, University of Toronto, Political Behaviour Group, Toronto.
- *A Word-Count Based Classifier of Politicians in the Australian Federal Parliament (1901–2018)*
  - 4 September 2019, European Consortium for Political Research, General Conference, Wroclaw.
  - 22 June 2019, European Political Science Association, Annual Conference, Belfast.
  - 18 June 2019, Quantitative Text Analysis Workshop, Dublin.
- *The Effect of Elections and Prime Ministers on Discussion in the Australian Federal Parliament (1901—2018)*
  - 11 January 2019, Political Methodology Specialist Group, Annual Conference, Warwick.
  - 11 December 2018, Max Planck Institute for Demographic Research, Rostock.
  - 22 November 2018, University of Toronto, Political Behaviour Group, Toronto.
  - 25 October 2018, Australian National University, Research School of Economics, Canberra.
  - 25 October 2018, Parliamentary Library, Parliament of Australia, Canberra.
  - 24 October 2018, Australian National University, School of Politics and International Relations, Canberra.
- *Exploring Australia's Hansard (1901–2017)*
  - (Poster) 7-9 September, 2018 Economic History Association Meeting, Montreal.
- *A Surnames-Based Analysis of Tasmanian Social Mobility (1803–2015)*
  - 20 May 2016, UC Berkeley, Economic History Lunch, Berkeley.
  - 2 May 2016, Australian National University, Research School of Economics, Canberra.