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).
- Co-lead, Data Sciences Institute Thematic Program in Reproducibility   November 2021 - *Current*
- 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*
- Faculty Affiliate, Schwartz Reisman Institute for Technology and Society   August 2020 - *Current*
- 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, *Telling Stories with Data*, (accepted for publication by CRC Press in May 2021).
  - Before we can research data science we need to know what it is. The central thesis of this book is that a revolution is needed in data science, and I propose a view of what it could look like. The book teaches the entire data science life cycle from developing research questions, data gathering and cleaning, analysis, and communication.

Edited books

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

Peer-reviewed articles

- Collins, Annie, and Rohan Alexander, 2022, 'Reproducibility of COVID-19 pre-prints', *Scientometrics*, 4 July, https://doi.org/10.1007/s11192-022-04418-2.
  - Abstract: To examine the reproducibility of COVID-19 research, we create a dataset of pre-prints posted to arXiv, bioRxiv, and medRxiv between 28 January 2020 and 30 June 2021 that are related to COVID-19. We extract the text from these pre-prints and parse them looking for keyword markers signaling the availability of the data and code underpinning the pre-print. For the pre-prints that are in our sample, we are unable to find markers of either open data or open code for 75% of those on arXiv, 67% of those on bioRxiv, and 79% of those on medRxiv.
  - Preprint version: https://arxiv.org/abs/2107.10724.
- 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.

Pre-prints

- 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/.

Software

- Alexander, Rohan, and Paul A. Hodgetts, 2021, `AustralianPoliticians`.
  - Provides access to biographical and political data about Australian federal politicians who served between 1901 and 2021. Published on CRAN and downloaded 1,742 times as at 27 April 2022.
- Hodgetts, Paul A., and Rohan Alexander, 2021, `cesR`.
  - Provides access to the Canadian Election Study Datasets. Published on CRAN and downloaded 1,585 times as at 27 April 2022.
- Alexander, Rohan, and A Mahfouz, 2021, `heapsofpapers`.
  - Makes it easier to respectfully download papers, CSVs, websites, and similar. Published on CRAN and downloaded 2,804 times as at 27 April 2022.

Grants

- 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 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

- *(Upcoming)* Instructor, 'Applied Statistics', Information Winter 2023
- *(Upcoming)* 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, 'INF2178: Experimental Design for Data Science', Information Winter 2021
- Instructor, 'STA304: Surveys, Sampling and Observational Data', Statistical Sciences Fall 2020
- Instructor, 'INF2178: Experimental Design for Data Science', Information Winter 2020

Supervisory experience

PhD

- Ciara Zogheib, Information September 2022 - *Current*

Masters

- Ciara Zogheib, Information (thesis supervisor) September 2021 - June 2022
  - Title: *The 'Paper and Pencil' Operations of Interdisciplinarity: Interdisciplinary Integration as Information Practice*
- Elisa Du, Applied Computing (academic supervisor) March 2021 - December 2021
  - Title: *Enriched Understanding of Retail Receipts for Personalized Financial Insights*

Undergraduate

- Callie Moore, Engineering (senior thesis) September 2021 - April 2022
  - Title: *Dyadic Representation in Canadian Parliament*
- Jeremy Hunt, Engineering (senior thesis) September 2020 - April 2021
  - Title: *Study of Software Tool Usage in Academics*

Service

Outreach

- Co-organizer, *Toronto Data Workshop* September 2019 - *Current*
- Organizer, *Florence Nightingale Day* 9 April 2022

– Promotes gender diversity in statistics and data science by providing high school students with teaching activities, mentoring sessions, and panel discussions with statisticians and data scientists.
- Organizer, *Toronto Workshop on Reproducibility*   25-26 February 2021 & 23-25 February 2022
  – Annual conference with three tracks: evaluating reproducibility, practices of reproducibility, and teaching reproducibility. Hosted online, and attracts global presenters and attendees.
- Co-organizer, *MRP Working Group*   January 2020 - November 2021

Departmental and University Committees

- Member, Search Committee, Dalla Lana School of Public Health, Assistant Professor—Data Science Methods for Population Health and Health Systems,   July 2022 - *Current*
- Chair, Data Sciences Speaker Series Selection Committee, Data Sciences Institute March 2022 - *Current*
- Member, Faculty of Arts and Sciences Computation and Data Science Education Working Group September 2021 - *Current*
- Member, BI Recruitment and Admission Committee, Information   July 2021 - *Current*
- Member, Graduate Committee, Statistical Sciences   July 2020 - *Current*
- 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
- Co-organiser, Applied Research and Education Seminar Series, Statistical Sciences July 2020 - *Current*
- Member, Search Committee, Information, Learning Hub Director and Teaching Stream Assistant Professor   November 2020 - April 2020
- Member, Awards Committee, Information   July 2020 - November 2020

Exam and Thesis Committees

- Chair, Statistical Sciences, PhD Oral Exam Committee, Michaël Lalancette   14 July 2022
- Member, Information, PhD Thesis Committee, Mirai Gendi   January 2022 - *Current*
  – Title: *Data science and process mining with application to digital forensics*
- 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

Editorial Service

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

Refereeing

- 2022: Journal of Statistics and Data Science Education (x2)
- 2021: CRC Press (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.

Other Service

- Advisor, *Statistical Sciences PhD Student Research Day*   25 May 2022
- Co-organizer, *Statistical Sciences PhD Practice Job Talk Day*   17 December 2021

Presentations

- *Searching for internal consistency with GPT-3*
  – *(Upcoming)* 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).
- *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).
- *Turning Our World Into Data*
  - – 7 July 2021, Harvard Biostatistics Data Science in Action Summer Camp (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.