

Data Staff Scientist  
Department of Biostatistics  
Fred Hutchinson Cancer Research Center

Updated: October 10, 2023  
Email: sean@seankross.com  
www.seankross.com

## RESEARCH INTERESTS

Human-computer interaction, data science, online learning, cybersecurity

## ACADEMIC POSITIONS

- 07/2022 – **The Fred Hutchinson Cancer Research Center**, Seattle, WA  
Data Staff Scientist
- 01/2018 – **The Johns Hopkins Bloomberg School of Public Health**, Baltimore, MD  
Affiliate Faculty of Biostatistics
- 01/2017 – 05/2017 **The University of Pennsylvania School of Design**, Philadelphia, PA  
Adjunct Faculty of Design

## EDUCATION

- 09/2017 – 06/2022 **The University of California, San Diego**, La Jolla, CA  
Ph.D. in Cognitive Science  
Advisor: Philip Guo
- 09/2017 – 03/2020 **The University of California, San Diego**, La Jolla, CA  
M.S. in Cognitive Science  
Advisor: Philip Guo
- 06/2012 – 05/2015 **The University of Maryland, College Park**, College Park, MD  
Bachelor of Science in Computer Science
- 09/2008 – 05/2012 **New York University**, New York, NY  
Bachelor of Arts in Biology

## AWARDS AND HONORS

- 08/2020 USENIX Security Distinguished Paper Award [C.10]
- 05/2019 CHI Honorable Mention Paper Award [C.7]
- 05/2012 New York University President's Service Award

## PRIOR EMPLOYMENT

- 05/2021 – 08/2021 **Microsoft Research**, Seattle, WA  
Research Intern – Research in Software Engineering (RiSE) Group  
Advised by Thomas Ball
- 05/2020 – 08/2020 **Microsoft Research**, New York, NY  
Research Intern – Computational Social Science Group  
Advised by Jake Hofman & Daniel Goldstein
- 09/2015 – 09/2017 **The Johns Hopkins Bloomberg School of Public Health**, Baltimore, MD  
Chief Technology Officer – Johns Hopkins Data Science Lab

- 01/2014 – 07/2015 **University of Maryland**, College Park, MD  
Research Assistant – Computational Biology Group – Advisor: Mihai Pop
- 01/2012 – 05/2012 **New York University**, New York, NY  
Research Assistant – Genomics Group – Advisor: Patrick Eichenberger

## PUBLICATIONS

### CONFERENCE PAPERS

- C.15 Sam Lau\*, **Sean Kross\***, Eugene Wu, Philip J. Guo (\*equal contribution). Teaching Data Science by Visualizing Data Table Transformations: Pandas Tutor for Python, Tidy Data Tutor for R, and SQL Tutor. In Proceedings of DataEd 2023: *ACM International Workshop on Data Systems Education*, June 2023. Seattle, WA, USA.
- C.14 **Sean Kross** and Philip J. Guo. Five Pedagogical Principles of a User-Centered Design Course that Prepares Computing Undergraduates for Industry Jobs. In Proceedings of SIGCSE 2022: *ACM Technical Symposium on Computer Science Education*, March 2022. Providence, RI, USA.
- C.13 **Sean Kross** and Philip J. Guo. Orienting, Framing, Bridging, Magic, and Counseling: How Data Scientists Navigate the Outer Loop of Client Collaborations in Industry and Academia. In Proceedings of CSCW 2021: *ACM Conference on Computer Supported Cooperative Work*, November 2021. Toronto, Canada.
- C.12 **Sean Kross**, Eszter Hargittai, Elissa M. Redmiles. Characterizing the Online Learning Landscape: What and How People Learn Online. In Proceedings of CSCW 2021: *ACM Conference on Computer Supported Cooperative Work*, November 2021. Toronto, Canada.
- C.11 Xiaoying Pu, **Sean Kross**, Jake Hofman, Daniel Goldstein. Datamations: Animated Explanations of Data Analysis Pipelines. In Proceedings of CHI 2021: *ACM Conference on Human Factors in Computing Systems*, May 2021. Yokohama, Japan.
- C.10 Elissa M. Redmiles, Noel Warford, Amritha Jayanti, Aravind Koneru, **Sean Kross**, Miraida Morales, Rock Stevens, Michelle L. Mazurek. A Comprehensive Quality Evaluation of Security and Privacy Advice on the Web. In Proceedings of the 2020 USENIX Security Symposium. August 2020. (**Distinguished Paper Award**)
- C.9 Elissa M. Redmiles, Lisa Maszkiewicz, Emily Hwang, Dhruv Kuchhal, Everest Liu, Miraida Morales, Denis Peskov, Sudha Rao, Rock Stevens, Kristina Gligorić, **Sean Kross**, Michelle L. Mazurek, Hal Daumé III. Comparing and Developing Tools to Measure the Readability of Domain-Specific Texts. In Proceedings of EMNLP 2019: *ACL SIGDAT Conference on Empirical Methods in Natural Language Processing*, November 2019. Hong Kong, China.
- C.8 **Sean Kross** and Philip J. Guo. End-User Programmers Repurposing End-User Programming Tools to Foster Diversity in Adult End-User Programming Education. In Proceedings of VL/HCC 2019: *IEEE Symposium on Visual Languages and Human-Centric Computing*, October 2019. Memphis, TN, USA.
- C.7 **Sean Kross** and Philip J. Guo. Practitioners Teaching Data Science in Industry and Academia: Expectations, Workflows, and Challenges. In Proceedings of CHI 2019: *ACM Conference on Human Factors in Computing Systems*, May 2019. Glasgow, UK. (**Honorable Mention Paper Award**)

- C.6 Elissa M. Redmiles, **Sean Kross**, Michelle L. Mazurek. How Well Do My Results Generalize? Comparing Security and Privacy Survey Results from MTurk, Web, and Telephone Samples. In Proceedings of S&P 2019: *IEEE Symposium on Security and Privacy*, May 2019. San Francisco, CA, USA.
- C.5 Elissa M. Redmiles, Ziyun Zhu, **Sean Kross**, Dhruv Kuchhal, Tudor Dumitras, Michelle L. Mazurek. Asking for a Friend: Evaluating Response Biases in Security User Studies. In Proceedings of CCS 2018: *ACM Conference on Computer and Communications Security*, October 2018. Toronto, ON, Canada.
- C.4 **Sean Kross** and Philip J. Guo. Students, Systems, and Interactions: Synthesizing the First Four Years of Learning@Scale and Charting the Future. In Proceedings of L@S 2018: *ACM Conference on Learning at Scale*, June 2018. London, UK.
- C.3 Elissa M. Redmiles, **Sean Kross**, Michelle L. Mazurek. Where is the Digital Divide? Examining the Impact of Socioeconomics on Security and Privacy Outcomes. In Proceedings of CHI 2017: *ACM Conference on Human Factors in Computing Systems*, May 2017. Denver, CO, USA.
- C.2 Elissa M. Redmiles, **Sean Kross**, Michelle L. Mazurek. How I Learned to be Secure: a Census-Representative Survey of Security Advice Sources and Behavior. In Proceedings of CCS 2016: *ACM Conference on Computer and Communications Security*, October 2016. Vienna, Austria.
- C.1 Elissa Redmiles, Mary Allison Abad, Isabella Coronado, **Sean Kross**, Amelia Malone. A Classroom Tested Accessible Multimedia Resource for Engaging Underrepresented Students in Computing. In Proceedings of SIGCSE 2015: *ACM Conference on Computer Science Education*, March 2015. Kansas City, MO, USA.

#### JOURNAL PAPERS

- J.4 Emily J. Winokur, Cherry Song, Estelita S. Leija, Joanne Chen, **Sean Kross**, Danielle Shamam, Marcelo Aguilar-Rivera, Laleh K. Quinn, Federico Rossano, Andrea A. Chiba. Reciprocity in Dyads and Triads: Female Rats Alter Their Prosocial Behavior According to the Social Context. In *Animal Behavior and Cognition*, 2023.
- J.3 **Sean Kross**, Jeffery T. Leek, John Muschelli. Ari: The Automated R Instructor. In *The R Journal*, 2020.
- J.2 Lucy D’Agostino McGowan, **Sean Kross**, Jeffery T. Leek. Tools for Analyzing R Code the Tidy Way. In *The R Journal*, 2020.
- J.1 **Sean Kross**, Roger D. Peng, Brian S. Caffo, Ira Gooding, Jeffery T. Leek. The Democratization of Data Science Education. In *The American Statistician*, 2020.

#### BOOKS

- B.2 **Sean Kross**. The Unix Workbench. Leanpub, 2017.
- B.1 Roger D. Peng, **Sean Kross**, Brooke Anderson. Mastering Software Development in R. Leanpub, 2017.

#### FUNDING

- F.2 Alfred P. Sloan Foundation. “Investigating How Domain Scientists and Computational Scientists Work Together in Software-driven Research Labs” \$49,681. PI: Philip Guo, co-PI: **Sean Kross**. 2021-2022.

F.1 Facebook Research: Economic Opportunity and Digital Platforms. “Learning Online: Who Benefits, Who is Left Behind?” \$50,000. PI: Eszter Hargittai, co-PI: **Sean Kross**, co-PI: Elissa M. Redmiles. 2020.

## TALKS

- Visualizing Data Analysis Pipelines with Pandas Tutor and Tidy Data Tutor. *Posit Conf*, September 2023.
- How Data Scientists Navigate the Outer Loop of Client Collaborations in Industry and Academia. *Joint Statistical Meetings*, August 2021.
- Create a Personal Website in 5 Minutes with Postcards. *Cascadia R Conference*, June 2021.
- The Challenges of Analytic Workflows: Perspectives from Data Science Educators. *The Eastern North American Region of the International Biometric Society (ENAR)*, Nashville, TN, March 2020.
- The Lean Course: Open and Collaborative Online Course Development. *Joint Statistical Meetings*, Vancouver, BC, Canada, July 2018.
- Lessons from Teaching Data Science to Over a Million People. *Crunch Conference*, Budapest, Hungary, October 2017.
- Beyond Axes: Simulating Systems with Interactive Graphics. *Joint Statistical Meetings*, Baltimore, MD, August 2017.
- Re-thinking the Value Added by a Data Repository. *Dataverse Community Meeting*, Cambridge, MA, June 2017.
- Scaling Data Science Education. *Seminário Internacional de Estatística*, Rio de Janeiro, Brazil, May 2016.

## TEACHING

### INSTRUCTOR

- UCSD COGS 127: Data-Driven User Experience and Product Design (Winter 2022)
- Johns Hopkins University on Coursera.org: The Unix Workbench (Continually since Summer 2017)
- The University of Pennsylvania MUSA 850: Data Wrangling and Data Visualization (Spring 2017)

### TEACHING ASSISTANT

- UCSD COGS 127: Data-Driven User Experience and Product Design (Fall 2019, 2020)
- UCSD COGS 9: Introduction to Data Science (Winter 2019)
- UCSD COGS 120/CSE 170: Human-Computer Interaction Design (Fall 2017, 2018)
- UCSD COGS 121 Human-Computer Interaction Programming Studio (Spring 2018)