STEVEN KERR

13347 Berlin | Germany
+49 152 28183817 | steven.miles.k@gmail.com
github.com/smkerr | website: smkerr.github.io

EDUCATION

M.Sc. Data Science for Public Policy **Sep 2021—Jun 2024**

The Hertie School of Governance, Berlin, Germany

- Interdisciplinary program focusing on applying data science to policy questions
- Receiving merit-based partial tuition waiver
- Current Grade: 1.6 (scale 1-6, 1 = best)

B.A. Economics, *with honours* **Aug 2015—Apr 2019**

Pepperdine University, Los Angeles, California, USA

- Thesis: "Effects of Foreign Exchange Rates on US-Japan Trade Relations"
- Final Grade: 3.68/A- (scale 0-4, 4 = best)

PROFESSIONAL EXPERIENCE

Student Researcher, Mercator Research Institute (MCC) **Jan 2024—Present**

Berlin, Germany

- Working with the Applied Sustainability Science Group to enhance evidence synthesis through the development of a living map of scientific literature on climate change and human health
- Developing both supervised and unsupervised machine learning models to correctly identify and categorize scientific literature on climate change and human health using Python
- Conducting citation network analysis to map the interconnectivity of scientific publications on climate change and human health

Teaching Assistant, Hertie School **Aug 2023—Jan 2024**

Berlin, Germany

- Led weekly lab sessions for the Introduction to Data Science course to demonstrate hands-on application of course concepts
- Educated students on key data science tools and programming languages, such as GitHub and R
- Prepared course materials detailing best practices for collecting, cleaning, and analyzing data
- Achieved 94% satisfaction rating from student course evaluations

Junior Data Scientist, CPC Analytics **Jun 2022—May 2023**

Berlin, Germany

- Assisted the WHO Pandemic Hub in developing the 2022-23 Mpox (Monkeypox) Outbreak: Global Trends Shiny application using R to inform both WHO decision-makers and the wider public
- Assisted in the development of an R-based tool using PubMed data that enables users to perform automated bibliometric analyses of health-related topics
- Analysed OECD data to examine trends in donor states' health-related official development assistance to developing countries

Procurement Analyst, U.S. Space Force **Aug 2019—Jul 2021**

Los Angeles, California, USA

- Procured supplies and services necessary to support $800M U.S. satellite launch program
- Generated time series forecasts, conducted industry analyses, and frequently drafted reports presenting decision-makers with viable options to support negotiations

Director of Education, Pepperdine Model United Nations **Apr 2018—Apr 2019**

Los Angeles, California, USA

- Led weekly 1hr 30min sessions to train 30+ undergraduate and graduate students in research methods, policy analysis, memo writing, public speaking and debate skills
- Formulated reports on a diverse set of topics such as climate-induced displacement, financing for renewable energy projects and the impact of climate change on food security

Economics Tutor, Pepperdine University **Aug 2018—Dec 2018**

Los Angeles, California, USA

- Demonstrated knowledge of advanced economic theory while tutoring 15 undergraduate students
- Created individual progress plans and organized group review sessions ahead of exams

Research Assistant, Pepperdine University **May 2017—Jun 2017**

Los Angeles, California, USA

- Conducted qualitative study examining student perspectives on diversity at Pepperdine University
- Performed literature review, ran 12 focus groups, and transcribed student interviews
- Generated report synthesising findings and presented key takeaways at research symposium

Associate Editor, Pepperdine University **Jan 2016—Apr 2017**

Los Angeles, California, USA

- Evaluated submissions for the *Pepperdine Journal of Communication Research*
- Nominated submissions based on academic rigour, organizational structure, and grammar

PUBLICATIONS

Mpox in Children and Adolescents during Multicountry Outbreak, 2022-2023 **Oct 2023**

Hoxha A, **Kerr SM**, Laurenson-Schafer H, Sklenovská N, Mirembe B, Nezu I, et al. Mpox in Children and Adolescents during Multicountry Outbreak, 2022–2023. Emerging Infectious Diseases. 2023 Oct ; 29(10):2125-2129. doi:10.3201/eid2910.230516.

Description of the First Global Outbreak of Mpox **Jul 2023**

Laurenson-Schafer H, Sklenovská N, Hoxha A, **Kerr SM**, Ndumbi P, Fitzner J, et al. Description of the first global outbreak of mpox: an analysis of global surveillance data. The Lancet Global Health. 2023 Jul; 11(7):e1012-e1023. doi:10.1016/S2214-109X(23)00198-5.

RESEARCH PROJECTS

Tutorial: Retrieval Augmented Generation with Citation **Dec 2023**

Deep Learning (GRAD-E1394)

- Developed a **Python** tutorial on how to implement a Retrieval Augmented Generation with Citation (RAG+C) chatbot to address the challenge of knowledge management in government
- RAG+C provides Large Language Models (LLMs) with additional context sourced from an external database which significantly improves response accuracy and avoids hallucinations
- The tutorial demonstrates fundamental RAG+C concepts and deploys a chatbot on Hugging Face

EU Procurement Network Analysis **Dec 2023**

Applied Network Analysis (GRAD-E1426)

- Examined network dynamics within EU public procurement markets based on a dataset of 234 public procurement markets across 26 European countries from 2008-2016 using **Python**
- Networks comprised buyers and suppliers of public contracts
- Investigated network characteristics associated with perceived government corruption

Israel-Hamas War News Coverage Text Analysis **Dec 2023**

Text as Data (GRAD-E1282)

- Investigated how news coverage of Israel-Hamas War evolved over time and the extent to which news coverage varied by news source by conducting a text analysis with **R**
- Sourced 4,000+ articles from various news sources including the New York Times, Al Jazeera, and Die Welt published in the weeks immediately preceding and following Hamas's October 7 attack
- Employed topic modeling and sentiment analysis to analyze trends in news coverage by group

COVID Fake News Detection **May 2022**

Machine Learning (GRAD-C24)

- Trained a machine learning algorithm to detect fake Tweets related to COVID-19 using **Python**
- Tested the accuracy of the model across time periods to measure the extent to which the effectiveness of such models degrade over time due to the rapidly evolving nature of fake news

UK COVID Dashboard Disaggregated by Age Group **Dec 2021**

Introduction to Data Science (GRAD-C11)

- Utilized **R** to create a Shiny app-hosted dashboard to track COVID case rates in England with added functionality of disaggregating case rates by age group
- Accessed UK government APIs to obtain COVID case rate data and geodata

Validating COVID Self-Tests with Image Recognition **Dec 2021**

Data Structures & Algorithms (GRAD-C9)

- Developed an algorithm capable of validating COVID self-tests using image recognition

- Used data from the MNIST database of handwritten digits, a collection of more than 70,000 examples, to train the Machine Learning algorithm to recognize handwritten serial numbers and COVID test results using **Python**

Effects of Foreign Exchange Rates on U.S.-Japan Trade Relations **Apr 2019**

Applied Econometrics (ECON 410)

- Utilized **R** to examine the extent to which fluctuations in the U.S. dollar-Japanese yen (USD-JPY) exchange rate impact bilateral trade between the U.S. and Japan
- Performed Granger causality tests, impulse-response tests, and vector autoregression (VAR) analyses to test hypothesis using statistical packages in R

TECHNICAL COMPETENCIES

R

Passed LinkedIn Skill Assessment for R (Top 5%)

- Led lab sessions for the Hertie School's Introduction to Data Science course where R served as the programming language of instruction
- Frequently use R to clean, wrangle, model, and visualize data for projects and reports
- Advanced familiarity with tidverse (e.g., tidyr, dplyr, ggplot), sf (geocoding), and shiny packages

Python

- Utilize Python to prepare data, train ML models, and visualize results in professional settings
- Advanced experience with pandas, numpy, sklearn, matplotlib and seaborn packages

SQL

Certificate in Using Databases with Python

- Experience with creating, retrieving, updating, and deleting database information with SQL
- Used SQL to create a database of web-scraped geo-location data for data visualization project

Machine Learning

- Fine-tuned an open-source Large Language Model (LLM) on corpus of nearly 3,500 technical documents to build a chatbot capable of answering technical questions while citing its sources
- Developed a ML algorithm capable of validating COVID-19 rapid self-tests using image recognition
- Utilized ML and Natural Language Processing to detect fake social media posts related to public health topics

Other Coding Skills

- Git/GitHub/GitLab
- Jupyter Notebook/R Markdown/Quarto
- Command Line/Terminal/Bash/Zsh
- Shiny Apps/Streamlit/Dashboards
- Web-scraping (HTML/XML)

- API Requests
- LaTeX

MS Office

- Advanced experience with using Excel formulas, functions, and pivot tables
- Deep familiarity formatting reports, incorporate graphs & charts, and create mock-ups in Word
- Routinely deliver presentations, trainings, and briefings using PowerPoint

HONOURS & AWARDS

- Merit-based Tuition Waiver, Hertie School **2021—2024**
- Regents' Scholar Award (Top Academic Scholarship), Pepperdine University **2015—2019**
- Cum Laude, Pepperdine University **2019**
- Best Delegate Award, National Model United Nations (China) **2018**
- Position Paper Award, National Model United Nations (New York) **2018**
- Blanche E. Seaver Faculty-Staff Scholarship, Pepperdine University **2018**
- Dean's List, Pepperdine University **2017**

COMMUNITY INVOLVEMENT

- Tech Society, Hertie School **2022—Present**
- Data Science Society, Hertie School **2021—Present**
- Sigma Iota Rho (ΣIP) National Honors Society **2017—2019**
- Student Government Representative, Pepperdine University **2016—2017**

LANGUAGES

English	**Spanish**	**German**	**French**
Native	B2	B1	B1