Steven Kerr

Data Scientist at CPC Analytics, Data Science for Public Policy (MDS) Graduate, Hertie School

Berlin, Germany steven.miles.k@gmail.com smkerr.github.io
smkerr linkedin.com/in/stevenmileskerr

Education

2021-2024 **MSc in Data Science for Public Policy**
Hertie School, Berlin, Germany
Thesis: From Clicks to Cases: Leveraging Wikipedia Pageviews to Predict Mpox Cases in the United States (winner, Master's Thesis Poster Competition 2024)
Grade: 1.5 (German scale from 1.0 to 5.0, 1.0 = best); merit-based partial tuition waiver

2015-2019 **BA in Economics, cum laude**
Pepperdine University, Los Angeles, California, USA
Thesis: Effects of Foreign Exchange Rates on US-Japan Trade Relations
Grade: 3.68 out of 4.0; Regents' Scholar Award (top academic scholarship)

Professional Training

07/2025-08/2025 **Outbreak Analytics and Applied Modelling in R** (short course)
London School of Hygiene & Tropical Medicine, Online

Professional Experience

09/2024-present **Data Scientist**
CPC Analytics, Berlin, Germany
Co-lead the six-person data science team of a data and policy consultancy. Clients include WHO offices such as the Hub for Pandemic and Epidemic Intelligence, the Health Emergencies Programme and the Regional Office for South-East Asia (SEARO), as well as the German NGO Welthungerhilfe.

- *LLM evidence tools:* Built an automated pipeline that tracks new research across bibliographic databases and preprint servers, including PubMed, OpenAlex and Crossref, and uses LLMs to extract evidence tied to source quotes. Applied it to speed up evidence synthesis for the 2026 Bundibugyo ebolavirus outbreak.
- *Mpox emergency response:* Maintained WHO's mpox dashboard, prepared analyses for five Emergency Committee meetings, co-organized WHO's Mpox Analytics Community of Practice, and contributed weekly nowcasts for the DRC.
- *Epidemiological parameters:* Member of the core development team of grEPI (WHO), shaping its data structure and access; co-author of the {epiparameter} R package.
- *Risk assessment:* Led development of the WHO Seasonal Risk Assessment Tool (Excel, R Shiny), published by WHO in 2026, and drafted the corresponding manuscript.
- *Public health intelligence:* Led requirements definition with SEARO staff for a system that turns open-source epidemic intelligence into daily briefings and bulletins.
- *Development finance:* R pipelines on health emergency finance for WHO. Led the German ODA analysis for Welthungerhilfe's *Kompass* (2025, 2026) and advised the authors on the arguments the data supports.

01/2024-07/2024 **Student Researcher**
Mercator Research Institute on Global Commons and Climate Change, Berlin, Germany
Built a living map of climate and health research, classifying literature with machine learning in Python and mapping citation networks.

08/2023-01/2024 **Teaching Assistant**
Hertie School, Berlin, Germany
Led weekly R and GitHub labs for Introduction to Data Science (94% satisfaction).

06/2022-05/2023 **Junior Data Scientist**
CPC Analytics, Berlin, Germany
Contributed to WHO's 2022-23 mpox dashboard, supported WHO Ukraine's winter risk assessment, and co-authored two papers with WHO.

08/2019-07/2021 **Procurement Analyst**
U.S. Space Force, Los Angeles, California, USA
Procured supplies and services for an $800M satellite launch program and built time series forecasts and industry analyses to inform contract negotiations.

Languages

English - native proficiency
German - full professional proficiency (C1)
Spanish - upper-intermediate proficiency (B2)
French - intermediate proficiency (B1)

Data Science Competencies

PROGRAMMING & SOFTWARE

Python (pandas, polars, scikit-learn, PyTorch, docling), R (tidyverse, Shiny, targets, sf, Epiverse), SQL, Git, Quarto, LaTeX, Docker, Claude Code

DATA ENGINEERING & CLOUD

AWS, Azure, Terraform, GitHub Actions, REST APIs, scheduled jobs, Databricks

DATA ANALYSIS & MODELING

Spatial analysis and regression, time series forecasting, machine learning, text and network analysis, composite indicators, outbreak analytics

LLMS

Anthropic and OpenAI APIs, Hugging Face, LangChain, RAG with citations, structured extraction (pydantic), embeddings and vector databases, automated output verification, fine-tuning

Publications

2023 **Peer-reviewed article, *Emerging Infectious Diseases***
Hoxha A, **Kerr SM**, et al. Mpox in Children and Adolescents during Multicountry Outbreak, 2022-2023.

2023 **Peer-reviewed article, *The Lancet Global Health***
Laurenson-Schafer H, Sklenovská N, Hoxha A, **Kerr SM**, et al. Description of the first global outbreak of mpox: an analysis of global surveillance data.

2026 **Preprint, *Research Square***
Lofungola A, Ajong B, Hoxha A, ..., **Kerr SM**, et al. The multifaceted mpox outbreaks in the Democratic Republic of the Congo, January 2024-August 2025.

Last update: November 24, 2026
Typeset in LaTeX