

# H. Long Nguyen



## Education

present  
|  
07/2021

### PhD in Sociology

Bielefeld University

Working title: The many facets of Twitter (data) from a spatial perspective

06/2021  
|  
10/2018

### MA in Social Policy

University of Bremen

Grade: 1.2

Thesis: Estimating heterogeneous effects of welfare retrenchments on government support using causal random forests (Grade: 1.0)

09/2018  
|  
04/2015

### BA in Social Science & Philosophy

Ruhr University Bochum

Grade: 1.1

Thesis: Contextual effects of socially segregated neighbourhoods on social participation in Dresden and Leipzig (Grade: 1.0)

## Work Experience

present  
|  
07/2021

### Research Fellow, Leibniz ScienceCampus SOEP RegioHub at Bielefeld University (SOEP@UBi)

Project 4: RegioData – Establishing an innovative research and teaching data infrastructure

present  
|  
06/2021

### Data Visualisation Freelancer

Static and interactive graphics and web applications for research communication

Latest clients: DeZIM (German Centre for Integration and Migration Research), Citizens For Europe e.V., Humboldt University of Berlin

06/2023  
|  
05/2023

### Visiting Researcher, German Federal Institute for Population Research (BiB)

Research area 2: Migration and Mobility

03/2021  
|  
05/2020

### Working Student, Data Warehouse and Business Intelligence

Data Management and Institutional Research Unit, University of Hamburg

12/2019  
|  
10/2019

### Intern, Market & Opinion Research

konkret – Institut für innovative Markt- & Meinungsforschung GmbH

✉ [long.nguyen@uni-bielefeld.de](mailto:long.nguyen@uni-bielefeld.de)

🏠 [lo-ng.netlify.app](https://lo-ng.netlify.app)

🐙 [github.com/long39ng](https://github.com/long39ng)

🐦 [twitter.com/long39ng](https://twitter.com/long39ng)

## Research interests

Computational social science, esp. text as data and spatial data

## Skills

Programming languages: R, Python, JavaScript/TypeScript, Rust

Data visualisation with R and JavaScript

Database management with SQL, esp. PostgreSQL

Containerisation with Docker

Version controlling with git; CI/CD with GitHub/GitLab

Markup languages: HTML, LaTeX, Markdown

## Languages

English (fluent), German (fluent), Vietnamese (native)

Last updated on 2023-11-30.

- 09/2019  
|  
04/2019 • **Student Assistant, Project MARDY: Modeling ARGumentation DYnamics in Political Discourse**  
SOCIUM Research Center on Inequality and Social Policy, University of Bremen
- 07/2019  
|  
01/2017 • **Student Assistant, Chair of Political Science I: State and Governance**  
Institute of Political Science, University of Hagen



## Research Articles

- 2023 • **X-GTA: The Cross-Topic German Twitter Archive**  
SocArXiv Preprint. DOI: [10.31235/osf.io/9tbd4](https://doi.org/10.31235/osf.io/9tbd4)  
Tsolak D, Knauff S, Kühne S, & **Nguyen HL**
- 2022 • **Efficient and Reliable Geocoding of German Twitter Data to Enable Spatial Data Linkage to Official Statistics and Other Data Sources**  
Frontiers in Sociology, 7. DOI: [10.3389/fsoc.2022.910111](https://doi.org/10.3389/fsoc.2022.910111)  
**Nguyen HL**, Tsolak D, Karmann A, Knauff S, & Kühne S



## Academic Peer Review

- 2022,  
2023 • **Transactions in GIS**



## Open-Source Software

- **R packages**
  - [badata](#): regional job market data from the German Employment Agency
  - [deairdata](#): R API client for air data from the German Environment Agency
  - [inkr](#): local access from R to all INKAR regional data
  - [listviewerlite](#): collapsible tree view of nested lists in R
  - [nutsdecoder](#): geocoding to NUTS region codes
  - [remmy](#): R API client for Lemmy
  - [restatis](#): R API client for the German Federal Statistical Office Database



## Summer Schools and Workshops

- 07/2022 • **Summer Institute in Computational Social Science (SICSS)**  
Oxford University



## Teaching

- Summer  
2023 • **Collection and Analysis of Digital Behavioural Data**  
MA Sociology – with Zaza Zindel
- Winter  
2022/23,  
Summer  
2022 • **Introduction to R**  
BSc Business Administration and Economics – with Jonas Bauer, Sebastian Büscher, Julia Dyck, Carlina Feldmann, Anna Karmann, and Lennart Oelschläger
- Summer  
2022 • **Geospatial Analysis with SOEP Data**  
PhD course at Bielefeld Graduate School in History and Sociology – with Simon Kühne
- Winter  
2021/22 • **Geospatial Data in the Social Sciences**  
PhD course at Bielefeld Graduate School in History and Sociology – with Simon Kühne

Winter  
2023/24,  
2021/22

- **Statistical Consulting**  
Project-based course in MSc Statistics/Data Science – with other project advisors

## Conference Presentations

- 11/2023 ● **Quetelet Seminar at UCLouvain: Unconventional Data Sources for Population Studies**  
Estimating internal migration flows using origin–destination data based on Twitter user profile locations
- 07/2023 ● **Conference of the European Survey Research Association (ESRA)**  
Estimating internal migration flows using origin–destination data based on Twitter user profile locations
- 07/2023 ● **Frontiers in Policy Relevant Population Research, German Federal Institute for Population Research (BiB)**  
Monitoring internal migration flows using origin–destination data based on Twitter user profile locations
- 03/2023 ● **Current Perspectives on Spatial Mobilities, German Institute for Employment Research (IAB)**  
Monitoring internal migration flows using origin–destination data based on Twitter user profile locations
- 02/2023 ● **Conference of Young Demographers**  
Monitoring internal migration flows using origin–destination data based on Twitter user profile locations
- 09/2022 ● **Congress of the German Sociological Association**  
Geokodierung deutscher Twitter-Daten für kleinräumige regionale Analysen
- 09/2022 ● **General Online Research**  
Geocoding Tweets Using Profile Locations: Incomplex Idea, Intricate Implementation
- 05/2022 ● **AAPOR Annual Conference**  
Efficient and Reliable Geocoding of Twitter Data to Enable Spatial Data Linkage with Official Statistics and Other Data Sources
- 03/2022 ● **Spring Conference, Section Methods of Empirical Social Research, German Sociological Association**  
Effiziente und reliable Geokodierung deutscher Twitter-Daten zur räumlichen Datenverknüpfung mit amtlichen Statistiken und anderen Datenquellen
- 09/2021 ● **General Online Research**  
The Discourse about Racism on German Social Media – A Big Data Analysis (co-presenter – with Anna Karmann, Dorian Tsolak, Stefan Knauff, Simon Kühne, Hendrik Lücking)

## Other Invited Talks

- 11/2021 ● **CorrelCon 2021**  
Creating your own themes and colour scales for ggplot2 ([slides](#))
- 03/2021 ● **Digital Social Summit – Strategien. Instrumente. Debatten.**  
Einführung in Data4Good (with Nina Hauser and Antonia Scherz)
- 11/2020 ● **CorrelCon 2020**  
Solving coding challenges for good (with speakers from other CorrelAid local chapters)

## Volunteering

- 04/2023 ● “Workshops for Ukraine” series (organised by Dariia Mykhailyshyna)  
Workshop: “Spatial Data Wrangling with R: A Comprehensive Guide”
- present  
|  
2019 ● Volunteer at CorrelAid
  - Volunteer in data science/software engineering projects for NPOs
  - Co-host of a bi-weekly data visualisation meetup

- present | 2017 Mentor for international students in the initiative Schulen: Partner der Zukunft (PASCH) by ZfA, PAD, and DAAD
- 2016 | 2015 Consultant for Higher Education Policy, General Students' Committee (AStA), University of Bochum; Deputy Chairman from 03/2016
- 2016 Campus guide for international students, International Office, University of Bochum



## Awards and Honours

- 11/2023 Travel and accommodation grant for the Quetelet Seminar at UCLouvain
- 07/2022 Travel grant for the Summer Institute in Computational Social Science (SICSS) at Oxford University
- 07/2022 Winner, "Top Secret Machine Learning Prediction Competition" at SICSS Oxford
- 03/2021 | 03/2019 Scholarship from a party political foundation
- 09/2018 | 10/2016 Scholarship from Ruhr University Bochum (Deutschlandstipendium)
- 07/2012 Scholarship from the Pädagogischer Austauschdienst (PAD) of the Kultusministerkonferenz