

# Teilnahmebescheinigung

**Malte Hückstädt**

hat von 14.–18. März 2022 an der Spring School

## **Webscraping und Data Mining mit Python**

unter der Leitung von

**Dr. Jan R. Riebling (Bergische Universität Wuppertal)**

teilgenommen

### **Inhalte des Workshops:**

#### **Computational Social Science und Einführung in Python (11,5 h)**

- *Einführung in die technische Infrastruktur Anaconda*
- *Vorstellung grundlegender Programmiertechniken und Vorgehensweisen im Rahmen der Computational Social Science*
- *Besprechung von Pythons Syntax, primitiven und weiterführenden Datentypen sowie relevanter Python-Module und Bibliotheken für die Forschung*

#### **APIs und Webscraping (7 h)**

- *Umgang mit webbasierten APIs (Application Programming Interfaces)*
- *Grundlegendes Vorgehen beim Webscraping, dem rekursiven Download von Webressourcen und der Extraktion von Informationen aus HTML-Dokumenten*

#### **Datenaufbereitung (7h)**

- *Aufbereitung von Daten und die Extraktion von Informationen aus Rohdaten, insbesondere Texten*
- *Arbeiten mit regulären Ausdrücken*

#### **Datenmanagement und Datenanalyse (4,5 h)**

- *Beurteilung der Datenqualität*
- *Techniken der Dokumentation und Archivierung von Daten*
- *Einführung in maschinelles Lernen*

Mainz, 18. März 2022

  
**Prof. Dr. Gunnar Otte**

Arbeitsbereich Sozialstrukturanalyse

Institut für Soziologie

Johannes Gutenberg-Universität Mainz

**Prof. Dr. Natascha Nisic**

Arbeitsbereich Soziologie und Methoden der  
quantitativen Sozialforschung

Institut für Soziologie

Johannes Gutenberg-Universität Mainz

