

# Indira Sen

COMPUTER SCIENCE P.H.D CANDIDATE INTERESTED IN THE APPLICATION OF NLP METHODS IN COMPUTATIONAL SOCIAL SCIENCE

Unter Sachsenhausen 6-8, Cologne 50667, Germany

+49(221)47694-253 | [indira.sen@rwth-aachen.de](mailto:indira.sen@rwth-aachen.de) | <https://indiiigo.github.io/> | [Indiiigo](#) | [@indiiigosky](#) | [Google Scholar](#)

## Research Interests

I'm interested in improving the measurement of psychosocial constructs, including political and prejudicial attitudes and behaviors, from digital trace data using mixed methods and natural language processing. My research aims to illuminate and characterize the potentials, pitfalls, and consequences of understanding such attitudes from unobtrusive digital traces.

## Education

### RWTH-Aachen and GESIS - Leibnitz Institute for Social Sciences

Cologne, Germany

P.H.D IN COMPUTER SCIENCE

Oct, 2018 - Oct, 2023 (exp)

**ADVISOR:** PROF. CLAUDIA WAGNER

### Indraprastha Institute of Information Technology (IIIT-D)

New Delhi, India

M.S IN COMPUTER SCIENCE

Aug, 2015 - Dec, 2017

**ADVISOR:** PROF. PONNURANGAM KUMARAGURU (PK)

**GPA:** 8.67/10

### Vellore Institute of Technology (VIT)

Vellore, India

B.S IN COMPUTER SCIENCE

Jul, 2010 - May, 2014

**GPA:** 9.1/10

## Honors & Awards

2022	<b>Research Award for Interdepartmental Collaboration</b> , for interdisciplinary research collaborations	GESIS
2021	<b>Summer Intern Award for Outstanding Innovation</b> , Nokia Bell Labs	Cambridge, UK / virtual
2021- 2023	<b>Best Reviewer Award</b> , ICWSM (2021, 2022, 2023), The WebConf (2023)	
2020	<b>Travel Grant for attending EMNLP</b> ,	virtual
2020	<b>Selected for SAGE Text-Mining Fellowship</b> , SAGE, Ocean	London, UK (virtual)
2020	<b>Selected as Facilitator for the Alan Turing Institute Data Study Group</b> , 2 week hackathon where we worked with an Industry partner (CatsAI) to deploy explainable AI predictions	virtual
2020	<b>Travel Grant</b> , For attending FAT* 2020	Barcelona, Spain
2019	<b>Travel Grant</b> , For attending Summer Institute in Computational Social Science (SICCS), Germany	Bamberg, Germany
2017	<b>TA Excellence Award</b> , Foundations of Computer Security	IIIT-Delhi, India
2017	<b>Accepted to Summer@EPFL</b> , One of 60 interns selected out of 2000 applications	EPFL, Switzerland
2015- 2016	<b>Dean's Merit List</b> , For Academic Performance	IIIT-Delhi, India
2015	<b>GATE Scholarship</b> , Scholarship for Master's Studies	
2011- 2014	<b>Merit Scholarship</b> , For Outstanding Academic Achievement	VIT, Vellore

## Publications

### Journal Articles

#### Insider Stories: Analyzing Internal Sustainability Efforts of Major US Companies from Online Reviews.

2023

INDIRA SEN, DANIELE QUERCIA, LICIA CAPRA, MATTEO MONTECCHI, AND SANJA ŠĆEPANOVIĆ.

HUMANITIES AND SOCIAL SCIENCES COMMUNICATIONS 10, NO. 1 (2023): 1-9.

**Depression at Work: Exploring Depression in Major US Companies from Online Reviews.** 2022

INDIRA SEN, DANIELE QUERCIA, MARIOS CONSTANTINIDES, MATTEO MONTECCHI, LICIA CAPRA, SANJA SCEPANOVIC, AND RENZO BIANCHI.

PROCEEDINGS OF THE ACM ON HUMAN-COMPUTER INTERACTION. TO BE PRESENTED AT THE 25TH ACM CONFERENCE ON COMPUTER-SUPPORTED COOPERATIVE WORK AND SOCIAL COMPUTING (CSCW)

**Visualizing Internal Sustainability Efforts in Big Companies.** 2022

CHIARA CECCARINI, EDYTA PAULINA BOGUCA, INDIRA SEN, MARIOS CONSTANTINIDES, CATIA PRANDI, AND DANIELE QUERCIA. IEEE COMPUTER GRAPHICS AND APPLICATIONS

**A Total Error Framework for Digital Traces of Human Behavior on Online Platforms** 2021

INDIRA SEN, FABIAN FLOECK, KATRIN WELLER, BERND WEISS, AND CLAUDIA WAGNER.

PUBLIC OPINION QUARTERLY

## Conference Articles

**Counterfactually Augmented Data and Unintended Bias: The Case of Sexism and Hate Speech Detection** 2022

INDIRA SEN, MATTIA SAMORY, CLAUDIA WAGNER AND ISABELLE AUGENSTEIN.

ANNUAL CONFERENCE OF THE NORTH AMERICAN CHAPTER OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS (NAACL) *short paper*

**How Does Counterfactually Augmented Data Impact Models for Social Computing Constructs?** 2021

INDIRA SEN, MATTIA SAMORY, FABIAN FLOECK, CLAUDIA WAGNER AND ISABELLE AUGENSTEIN.

CONFERENCE ON EMPIRICAL METHODS ON NATURAL LANGUAGE PROCESSING (EMNLP)

**“Call me sexist, but...” : Revisiting Sexism Detection Using Psychological Scales and Adversarial Samples** 2021

MATTIA SAMORY, INDIRA SEN, JULIAN KOHNE, FABIAN FLOECK, AND CLAUDIA WAGNER.

INTERNATIONAL CONFERENCE ON WEB AND SOCIAL MEDIA (ICWSM)

**On the Reliability and Validity of Detecting Approval of Political Actors in Tweets** 2020

INDIRA SEN, FABIAN FLOECK, AND CLAUDIA WAGNER.

CONFERENCE ON EMPIRICAL METHODS ON NATURAL LANGUAGE PROCESSING (EMNLP)

**Empowering First Responders through Automated Multimodal Content Moderation** 2018

DIVAM GUPTA, INDIRA SEN, NIHARIKA SACHDEVA, PONNURANGAM KUMARAGURU, ARUN BALAJI BUDURU.

IEEE INTERNATIONAL CONFERENCE ON COGNITIVE COMPUTING *Best Paper Award*

**Worth its Weight in Likes: Towards Detecting Fake Likes on Instagram** 2018

INDIRA SEN, ANUPAMA AGGARWAL, SHIVEN MIAN, SIDDHARTH SINGH, PONNURANGAM KUMARAGURU, ANWITAMAN DATTA.

WEBSCI *short paper*

## Workshop Articles

**What Human Rationales Tell us about Machine Explanations** 2021

INDIRA SEN, MATTIA SAMORY, FABIAN FLOECK, CLAUDIA WAGNER

HCI+NLP WORKSHOP, IN CONJUNCTION WITH EACL

**Measuring Public Opinion Online: Potentials and Pitfalls** 2020

INDIRA SEN, FABIAN FLOECK, CLAUDIA WAGNER

WOMEN IN MACHINE LEARNING (WIML), IN CONJUNCTION WITH NEURIPS *Poster*

**A Twitter Corpus for Hindi-English Code Mixed POS Tagging** 2018

KUSHAGRA SINGH, INDIRA SEN, PONNURANGAM KUMARAGURU.

WORKSHOP ON NATURAL LANGUAGE PROCESSING FOR SOCIAL MEDIA (SOCIALNLP), IN CONJUNCTION WITH ACL *Data Paper*

**Language Identification and Named Entity Recognition in Hinglish Code Mixed Tweets** 2018

KUSHAGRA SINGH, INDIRA SEN, PONNURANGAM KUMARAGURU. ACL STUDENT RESEARCH WORKSHOP

## Experience

**Snap Inc.** Santa Monica, USA

RESEARCH INTERN, COMPUTATIONAL SOCIAL SCIENCE. ADVISOR: DR. FRANCESCO BARBIERI

Jul 2022 - Oct 2022

- Developing a Stakeholder-centred Framework for Multilingual Sensitive Content Detection

## Nokia Bell Labs

INTERN AT SOCIAL DYNAMICS GROUP. ADVISOR: DR. DANIELE QUERCIA

Cambridge, UK (virtual)

Jun 2021 - Aug 2021

- Understanding and Characterizing the sustainability efforts of companies from employee reviews.
- Developed mixed-method techniques for detecting employees' views on their employer's sustainability efforts using a deep-learning framework (SBERT)
- Results from the internship were published at CSCW'22 and IEEE Vis'22.

## SAGE Ocean

SAGE OCEAN FELLOW. ADVISOR: DR. DANIELA DUCA

London, UK (virtual)

Oct 2020 - Dec 2020

- Contributing to a Text cleaning WebApp called Texti
- Working on NLP transformations in Python and documentation
- Results from the Fellowship are described here.

## Computational Social Science Department, GESIS

DOCTORAL RESEARCHER. ADVISORS: PROF. CLAUDIA WAGNER, DR. KATRIN WELLER

Cologne, Germany

Oct 2018 - Present

- Identifying, Characterizing and Quantifying errors that occur when measuring attitudes from Digital Traces
- Building robust and interpretable NLP methods for measuring social science phenomena
- Large scale Twitter data collection using Amazon Web Services (AWS) as a service for Social Scientists.

## Precog Research Lab, IIIT Delhi

RESEARCH ASSOCIATE. ADVISORS: PROF. PONNURANGAM KUMARAGURU AND PROF. ARUN BALAJI BUDURU

New Delhi, India

Dec 2017 - Aug 2018

- Landscaping the fake like marketplace on Instagram and characterizing fake likers, and their attack vectors.
- Detecting Multimodal Hatespeech on Social Media

## CHILI Lab, École polytechnique fédérale de Lausanne

SUMMER INTERN, ADVISOR: PROF. PIERRE DILLENBOURG

Lausanne, Switzerland

Jun 2017 - Aug 2017

- Worked on inferring dependencies between EPFL courses based on Concept Mining and grading patterns

## Nanyang Technological University

RESEARCH ASSISTANT, ADVISOR: PROF. ANWITAMAN DUTTA

Singapore

Jan 2017 - May 2017

- Master's Thesis: Worked on detecting fake engagement on Instagram.

# Teaching, Mentorship, and Academic Service

## Translation Tutorial on the Total Error Paradigm for Social Inferences

FAT\* 2020

Barcelona, Spain

Jan 2020

- Tutorial on diagnosing errors and biases in observational studies from digital traces

## Introduction to Python and Data Science with Python

UNIVERSITY OF COLOGNE

Cologne, Germany

Oct, 2019

- Interactive, hands-on five-day lecture series for teaching introductory python concepts to social science doctoral candidates

## Mentorship

SUPERVISED ON CO-SUPERVISED THESES

- Esi Doda (RWTH Aachen). Master's Thesis. Racial Disparities in Social Media Commentary about Football
- Vani Agarwal (IIIT-D). Master's Thesis. Improving Content Moderation through Co-Training
- Kushagra Singh (IIIT-D). Bachelor's Thesis. Information Extraction Tools for Code-Mixed Hindi-English Social Media Content

## PC member

WEBCONF (2023), ICWSM (2021, 2022), ACL STUDENT RESEARCH WORKSHOP (2020, 2021, 2022), WIKIWORKSHOP (2022),

HCI+NLP WORKSHOP (2022)

## Reviewer

CSCW (2019, 2020), THEWEBCONF (2020, 2021), EMNLP (2020), ACL ROLLING REVIEW (2021, 2022), EPJ DATA SCIENCE (2022)

# Skills

**Programming:** Python, R, Java, C++, C, MongoDB, SQL, Bash

**Libraries/Tools:** scikit-learn, pytorch, tensorflow, huggingface, keras, pandas, NLTK, NetworkX, Amazon Web Services, Amazon Mechanical Turk, Prolific, Qualtrics