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@gesis,org | 🏠 https://indiiigo.github.io/ | Indiiigo | 🐦 @indiiigosky | Google Scholar

Research Interests

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

Publications

Journal Articles

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 Bogucka, **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.   *short paper*
Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)

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.   *Best Paper Award*
IEEE International Conference on Cognitive Computing

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

Indira Sen, Anupama Aggarwal, Shiven Mian, Siddharth Singh, Ponnurangam Kumaraguru, Anwitaman Datta.   *short paper*
WebSci

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   *Poster*
Women in Machine Learning (WiML), In conjunction with Neurips

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

Kushagra Singh, **Indira Sen**, Ponnurangam Kumaraguru.   *Data Paper*
Workshop on Natural Language Processing for Social Media (SocialNLP), In conjunction with ACL

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   *Cambridge, UK (virtual)*
Intern at Social Dynamics Group. Advisor: Dr. Daniele Quercia   *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   *London, UK (virtual)*

SAGE Ocean Fellow. Advisor: Dr. Daniela Duca   *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   *Cologne, Germany*

Doctoral Researcher. Advisors: Prof. Claudia Wagner, Dr. Katrin Weller   *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   *New Delhi, India*

Research Associate. Advisors: Prof. Ponnurangam Kumaraguru and Prof. Arun Balaji Buduru   *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   *Lausanne, Switzerland*

Summer Intern, Advisor: Prof. Pierre Dillenbourg   *Jun 2017– Aug 2017*

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

Nanyang Technological University   *Singapore*

Research Assistant, Advisor: Prof. Anwitaman Dutta   *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   *Barcelona, Spain*

FAT* 2020   *Jan 2020*

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

Introduction to Python and Data Science with Python   *Cologne, Germany*

University of Cologne   *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