

# Sudipto Ghosh

---

CONTACT	<i>Phone:</i> +91-8826358310	<i>E-mail:</i> <a href="mailto:sudipto.ghosh@scai.iitd.ac.in">sudipto.ghosh@scai.iitd.ac.in</a>
KEY SKILLS	LLMs, Knowledge Graphs, Web Development, Data Mining, Data Analysis, Machine Learning, Project Management, Software Engineering, System Administration, Research Work.	
EDUCATION	<b>Indian Institute of Technology Delhi</b> , India Ph.D. in Artificial Intelligence, July 2024 – Present  <b>University of Delhi</b> , India M.Sc. in Computer Science, 2022 – 2024 B.Sc. (Hons.) in Computer Science, 2019 – 2022 <i>Recipient of University Gold Medal</i>	
PROFESSIONAL EXPERIENCE	<b>Indian Institute of Technology Delhi</b> <i>Doctoral Researcher · Full-Time</i> <b>July 2024 — Present</b> Working on Interpretability of LLMs and Knowledge Distillation at the Laboratory for Computational Social Systems (LCS2) under the supervision of Dr. Tanmoy Chakraborty.	
	<b>IBM India Research Labs</b> <i>Research Intern (GRM) · Internship</i> <b>February 2024 — June 2024</b> Worked on my M.Sc. major project under the IBM Global Remote Mentoring programme. Mentors: Balaji Ganesan (Senior Research Engineer, IBM Research), Prof. Vasudha Bhatnagar and Dr. Vikas Kumar (Department of Computer Science, University of Delhi).	
	<b>Jhatkaa.org</b> <i>Junior Software Developer · Part-Time</i> <b>August 2022 — May 2023</b> Worked with a RoR/PHP/Node.js/Django stack, developed and maintained various legacy action tools and open-source application deployments. Managed data, cloud infrastructure across providers and DevOps pipelines. Collaborated with external vendors and upstream developers.	
	<b>NPTEL</b> <i>Teaching Assistant</i> <b>January 2023 — March 2023</b> Facilitated the NOC: Introduction to Parallel Programming in OpenMP course run by IIT Delhi under Prof. Yogish Sabharwal, Adjunct Faculty, Dept. of CSE, IIT Delhi.	
	<b>iHub Anubhuti - IIITD Foundation</b> <i>Project Intern</i> <b>October 2021 — February 2022</b> Responsible for developing pipelines for Knowledge-based Question Answering in the legal domain under the guidance of Dr. Sarika Jain, Assistant Professor, Department of Computer Applications, National Institute of Technology Kurukshetra and Dr. Nandana Mihindukulasooriya, IBM Research, Dublin, Ireland.	
	<b>Major League Hacking</b> <i>Open Source Fellow</i> <b>June 2021 — August 2021</b> Worked on a JavaScript project under Jordan Harband (Member TC39, ECMA) for an npm package <i>can-merge</i> that polls the GitHub GraphQL API and allows custom scripts to run once all status checks and CI runs have completed. Participated in daily standups, pair programming, internal hackathons, CTF and code review sessions.	
	<b>StreetHack</b> , Jaymant and Sons OPC Pvt. Ltd., Noida, India <i>SDE Intern</i> <b>July 2019 — September 2019</b> Worked on a minimum viable product on the MERN stack. Developed RESTful APIs and a basic front-end. Employed OAuth2.0 and JWT authentication and dynamic resume generation.	

## PROJECTS

- **On Natural Language Generation with LLMs** — Generated datasets using LLaMA-2 and Falcon-LM and analyzed the generations for their lexical diversity, performed topic modelling and compared topic coherence, and compared sentiment intensities in the generations.
- **Distinguishing between AI-generated and Human-written Text: A Complex Network Approach** — Modelled text as a complex network and performed network analysis to extract features and experimented with models to classify classes of text apart.
- **DUCS Voting System** — Full-stack web application for conducting elections. Built on top of the Next.js framework. Was used in production for the DUCSS Elections in October 2023.
- **AQI Monitor** — Winning Submission at the [OutSystems, Vonage, LINKIT 2020 Developer Challenge Hackathon](#). Portal for monitoring air quality information and notification service on a range of communication channels including WhatsApp, Viber, Facebook Messenger and SMS.
- **stay.io** — Simple, scalable and containerized rating application for accommodation listings with role-based access control.
- **SQL2GeoJSON** — RESTful API service written in JavaScript that parses geometric data as stored in spatial databases (MySQL and PostgreSQL) and converts into GeoJSON for use with open-source mapping libraries like Leaflet.

## TECHNOLOGIES

- Languages: C, C++, Ruby, Python, R, Java, C#, JavaScript, PHP, SQL, HTML, CSS.
- Web Frameworks and Libraries: Ruby on Rails, Django, React, Node.js, Spring Boot.
- Databases: PostgreSQL, MySQL, IBM DB2, MongoDB, Cassandra, Elasticsearch, Redis.
- ML/DL Frameworks: Tensorflow, Keras, PyTorch, MLFlow, IBM Watson.
- NLP Tech: Protege, GATE, Stanza, spaCy, NLTK, Transformers.
- DevOps: CI/CD, Jenkins, Docker, Kubernetes.

## PUBLICATIONS

S. Ghosh, D. Verma, B. Ganesan, P. Bindal, V. Kumar, V. Bhatnagar. (August 2024). *InLegalLLaMA: Indian Legal Knowledge Enhanced Large Language Models*. Presented at the First OpenKG International Workshop, Large Knowledge-Enhanced Models at the 33rd International Joint Conference on Artificial Intelligence (IJCAI 2024), Jeju.

S. Ghosh, D. Verma, B. Ganesan, P. Bindal, V. Kumar, V. Bhatnagar. (March 2024). *Human Centered AI for Indian Legal Text Analysis*. Submitted to IJCAI 2024, Jeju. [arxiv.org/abs/2403.10944](https://arxiv.org/abs/2403.10944).

S. Jain, P. Harde, N. Mihindukulasooriya, S. Ghosh, et al. (May 2022). *Constructing a Knowledge Graph from Indian Legal Domain Corpus*. Presented at the First International Workshop on Knowledge Graph Generation From Text (TEXT2KG 2022) co-located with the 19th Extended Semantic Web Conference (ESWC 2022), Hersonissos, Greece, in CEUR Workshop Proceedings, Vol. 3184, Aug. 2022, pp. 80-93. [Online]. Available: [CEUR-WS.org/Vol-3184/TEXT2KG\\_Paper\\_6.pdf](https://ceur-ws.org/Vol-3184/TEXT2KG_Paper_6.pdf).

## ACHIEVEMENTS

**Qualified for Lecturership** – University Grants Commission, Dept. of Higher Education, GoI  
**University Gold Medal in B.Sc. (Hons.) Computer Science (2022)** – University of Delhi  
**Grand Prize Winner (2020)** — OutSystems, Vonage, LINKIT Developer Challenge Hackathon  
**INSPIRE Award MANAK (2015)** — Dept. of Science and Technology, Government of India

## VOLUNTEERING EXPERIENCE

**Department of Computer Science, University of Delhi**

*Treasurer, DUCSS; Coordinator, DUCS Research Club*

**October 2023 — June 2024**

**Microsoft Learn**

*Beta Student Ambassador*

**July 2021 — June 2022**

**Enactus, Atma Ram Sanatan Dharma College, New Delhi**

*President*

**July 2021 — May 2022**

*Deputy Research Head*

**July 2020 — July 2021**

*Executive Member*

**August 2019 — July 2020**

**Balvantray Mehta Vidya Bhawan ASMA, New Delhi**

*Science Captain*

**April 2017 — March 2018**

*President, Interact Club*

**September 2015 — February 2016**

REFERENCES

**Prof. Vasudha Bhatnagar**

*Professor, Department of Computer Science*  
Faculty of Mathematical Sciences  
University of Delhi, New Delhi, India  
*vbhatnagar@cs.du.ac.in*

**Dr. Vikas Kumar**

*Asst. Professor, Department of Computer Science*  
Faculty of Mathematical Sciences  
University of Delhi, New Delhi, India  
*vikas@cs.du.ac.in*