## *Kanoon Sarathi :*
A Multilingual Portal for the Indian Judicial System

The Background Story

The Digital India Mission launched by the Government of India in 2015 served as a beneficiary for many other schemes or initiatives and transformed the country as a whole. The legal industry is one of the major drivers for the stability of the society and the economy and thus it must be digitized. The major challenges observed in this process are unstructured and vague representation of the documents, legal jargon, inadequate signposting of key information, multiple interpretations of the information, only a subset of the entire Acts and Laws is relevant in context, and non-adaptability of the stakeholders for the digital shift. The COVID-19 pandemic had hit the world severely and accelerated the transition to the virtual world at an unprecedented rate. The court cases started to be heard through video conferencing, lawyers started doing online consultations, and the law firms started connecting to their clients through social media.

The Government of India has taken active efforts to present all laws/statutes along with their amendments at indiacode.nic.in and all court judgments at judis.nic.in. Some noteworthy developments of digitization and the application of AI to the Indian Judicial System are LIMBS, E-Courts, SUVAAS, SCI-Interact, SUPACE, SPOTDRAFT, CASEMINE, CASEIQ, NEARLAW, and PRACTICELEAGUE.
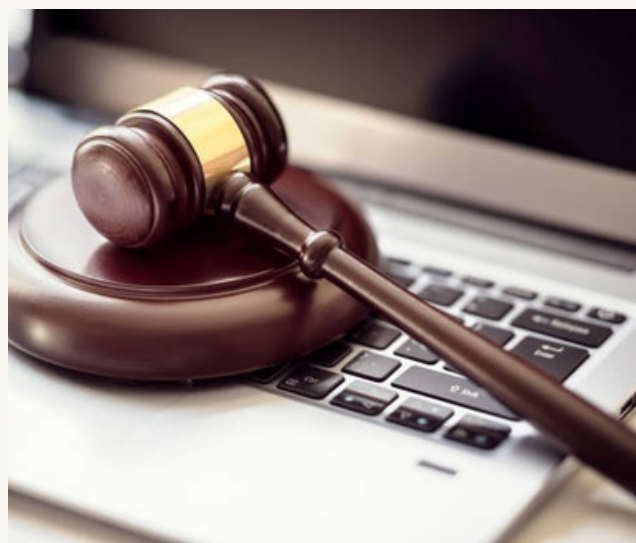
As of May 2022, over 4.7 crore cases are pending in different Indian courts (Data: The Hindu). Artificial Intelligence (AI) can go far beyond just automation in assisting the judges and lawyers. The theory of an AI agent (i.e., becoming better and better, the more it is practiced and the more data it is fed) matches with the way of working of a law system. AI-powered legal automation will be commonplace in the next five years to come. Artificial Intelligence can further aid the litigants, lawyers, and the judiciary through better document representation, document classification, discovering relevant documents, Online Dispute Resolution, Recovering Case Histories, Document Analysis, Legal Search, Strategizing the Case, Legal Summarization, Judgment Prediction, and much more.

The Team

Dr. Sarika Jain - Assistant Professor, National Institute of Technology, Kurukshetra is the Principal Investigator for this project. Her team includes one JRF, **Ms Pooja Harde** and several interns, namely Sudipto Ghosh, Dev Nirwal, Deepak Jaglan, Ankush Bisht, Abhinav Dubey, Tejas Mahajan, and Chirag Garg.

Kanoon Sarathi:

The two camps of AI, symbolic approaches (rule-based) and the sub-symbolic (statistical and machine learning) approaches, have their own say and their own pros and cons. The hybrid of the two camps, called the neuro-symbolic AI is the way forward to Artificial General Intelligence. Kanoon Sarathi is built upon this hybrid approach where the human cognition meets the standalone Statistical AI capabilities. Kanoon Sarathi is a multilingual portal for the Indian Judicial System that offers a platform for legal data curation, management, and sharing. This knowledge management platform will facilitate the interpretability and explainability of the decision support and the reasoning procedures applied

Currently Kanoon Sarthi focuses upon three use cases:

- Open data access for legal domain experts as well as to the common public to improve their involvement in understanding the judicial system in the form of Acts and Laws, Court Cases information, and so on
- Finding similar cases based upon the context present in the court case documents.
- Legal judgment prediction to help the judges decide the acceptance ratio of the appeal/petition with explainability.

Simply putting -

"We are helping various people to give relevant insight from their perspective into the law related documents that are generating today."

Dr. Sarika Jain

Milestones Achieved:

1. Presented and Under Publication with CEUR Workshop Proceedings (Scopus Indexed);
Title: Constructing a Knowledge Graph from Indian Legal Domain Corpus;
Venue: International Workshop On Knowledge Graph Generation From Text (Text2KG) Co-located with the Extended Semantic Web Conference (ESWC 2022);
Team: Sarika Jain, Sudipto Ghosh, Nandana Mihindukulasooriya, Abhinav Dubey, Pooja Harde, Ankush Bisht

2. Presented and Under Publication with Springer LNEE Series (Scopus Indexed);
Title: Investigating the Similarity of Court Decisions;
Venue: International Health Informatics Conference (IHIC 2022);
Team: Sarika Jain, Deepak Jaglan, Kapil Gupta

3. Presented and Under Publication with Springer LNEE Series (Scopus Indexed)
Title: NyOn: A Multilingual Modular Legal Ontology for Representing Court Judgements;
Venue: The International Semantic Intelligence Conference ISIC 2022;
Team: Sarika Jain, Pooja Harde, Nandana Mihindukulasooriya

Blog by Dr. Sarika Jain, Assistant Professor - National Institute of Technology, Kurukshetra

Research Update

The AI system that picks carcinogens out of the chemical crowd

Dr. Gaurav Ahuja and his team has developed a machine-learning technique to identify carcinogens from a set of six properties, such as ability to attract electrons, that are established for many carcinogens.

Dr. Gaurav Ahuja

Assistant Professor
Department of Computational
Biology, IIIT Delhi

The Research offers a new approach to accelerate detection of cancers. The team invented " Metabokiller" and used Metabokiller to virtually sift through 217, 921 metabolites stored in the Human Metabolome Database. Metabokiller is a promising new computer-aided method to detect carcinogenic human metabolites accurately and decipher how they trigger cancers.

The Team

| Dr. Debarka Sengupta | Ayushi Mittal | Sakshi Arora | Vishakha Gautam | Sanjay Kumar Mohanty |

EcTracker was the first product from TiH funding:
https://academic.oup.com/bib/article-abstract/22/6/bbab237/6309926

Most cancers are caused by exposure to carcinogens, compounds that make a cell malignant by altering its genome. Scientists typically have to use expensive, time-consuming animal tests to screen possible carcinogens. Machine learning could speed up the process, but earlier models were trained using a small number of known carcinogens and non-carcinogens, and cannot always predict the carcinogenicity of an unfamiliar compound.
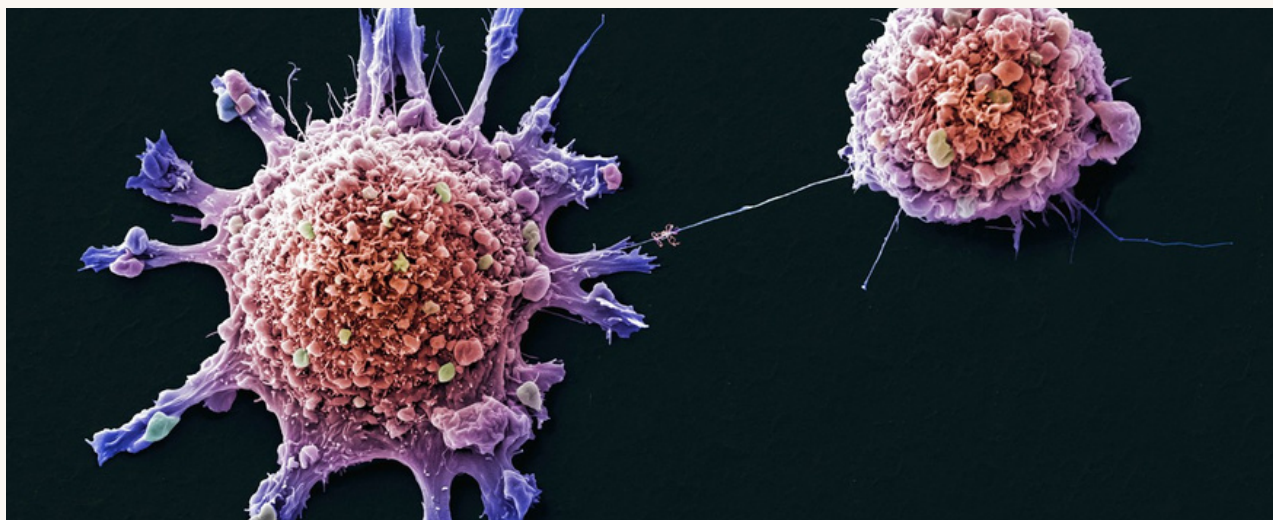
The researchers manually created 6 data sets, each with data on one property, encompassing a total of more than 35,000 compounds. They then trained and tested their models on these data sets. The software flagged two previously untested substances as carcinogens. Experiments on human and yeast cells showed that the compounds triggered changes associated with malignancy, validating the AI prediction.

The research work has constantly been supported by **iHub Anubhuti-IIITD Foundation**. **DBT & SERB** has also supported the research work. **Nature** and **Nature India** published the research on their respected pages.

Link to read the articles shared by nature.com and nature india :

https://www.nature.com/articles/d41586-022-02165-1

https://www.nature.com/articles/d44151-022-00084-8

It predicted that about 37.5% of the metabolites were carcinogenic. Of these, 38 had been identified previously. Some of the carcinogenic metabolites were melphalan, sulphur mustard, dichloroethane, chloral and trichloroethane.

The researchers validated the model's efficiency by exposing budding yeast cells and specific human lung epithelial cells to 4-nitrocatechol and 3,4-dihydroxyphenylacetic acid, two predicted carcinogenic metabolites. Both induced malignant changes in these cells.
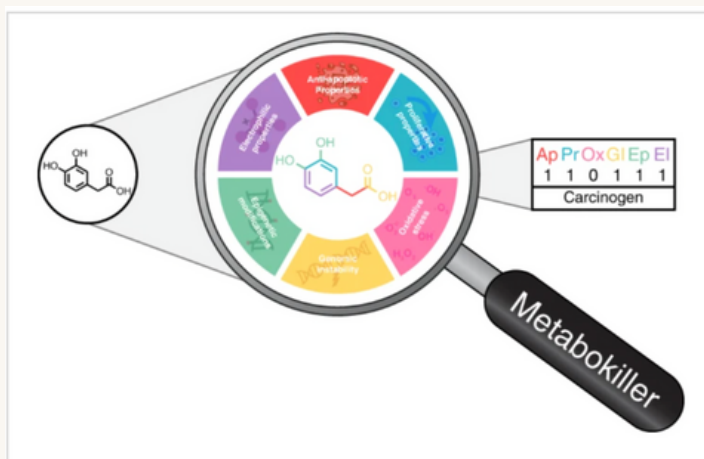
*Amrit Grand Challenge - जन*CARE

Call for proposals under Amrit Grand Challenge "जनCARE", Reimagining the Healthcare Delievery-Touching a billion lives was announced on 26th January 2022. The program was aligned with the National Digital Health Mission(NDHM) and would support 75 Healthcare innovations from start-ups & entrepreneurs in Digital Healthtech, Telemedicine, mHealth using Big Data, Blockchain, AI/ML & other technologies to strengthen the Healthcare ecosystem in India.

The program would provide extended support for funding, mentorship, and scaling up in several states, Medtech Industries, hospitals & corporates.

The funding will be provided under 3 categories:

- Ideation & Testing

- Pre-commercialization

- Multi-centric Product Deployment:

The evaluation process for Amrit Grand Challenge – The two-day online interview for stage 2 selection process was held on 3rd and 4th August 2022 in which 10 applicants were selected for the stage 3 evaluation process. The stage 3 evaluation process under the AGC जनCARE challenge was held on 5th August 2022 in Hyderabad.

The manager of **iHub Anubhuti-IIITD Foundation**, Mr. **Abhishek Bhattacharya**, partner-जनCARE has attended the evaluation process in person and was the part of selection committee.

BIRAC in partnership with iHub Anubhuti – IIITD Foundation and other partners will ensure that great ideas seamlessly reach the last mile. The latest collaboration between BIRAC and iHUB in digital health initiative i.e., जनCARE Challenge is a testament to the significant impact that can be brought together.

The official result is yet to be announced by BIRAC

Publications till Now

Year 2022

- Hai Q. Dinh, Bhanu Pratap Yadav, Sachin Pathak, Abhyendra Prasad, Ashish Kumar Upadhyay, Woraphon Yamaka, **Anuradha Sharma**, *$\mathbb{Z}4\mathbb{Z}4\mathbb{Z}4$-additive cyclic codes are asymptotically good.*

- Gopendra Vikram Singh, Mauajama Firdaus, Shambhavi, Shruti Mishra, **Asif Ekbal** , *Knowing What to Say: Towards knowledge grounded code-mixed response generation for open-domain conversations.*

- Gayatri Panda, Neha Mishra, Disha Sharma, Rintu Kutum, Rahul C. Bhoyar, Abhinav Jain, Mohamed Imran, Vigneshwar Senthilvel, Mohit Kumar Divakar, Anushree Mishra, Parth Garg, Priyanka Banerjee, Sridhar Sivasubbu, Vinod Scaria and **Arjun Ray** , *Comprehensive Assessment of Indian Variations in the Druggable Kinome Landscape Highlights Distinct Insights at the Sequence, Structure and Pharmacogenomic Stratum.*

- Asra Sakeen Wani, Divyanshu Kumar Singh, **Pushpendra Singh**, "*Hartal (Strike) Happens Here Everyday": Understanding Impact of Disruption on Education in Kashmir.*

- Deepika Yadav, Kirti Dabas, Prerna Malik, Anushka Bhandari, **Pushpendra Singh**, "*Should I visit the clinic*": *Analyzing WhatsApp-mediated Online Health Support for Expectant and New Mothers in Rural India.*

- Anupriya Tuli, Surbhi Singh, Rikita Narula, Neha Kumar, **Pushpendra Singh**, *Rethinking Menstrual Trackers Towards Period-Positive Ecologies.*

- **A V Subramanyam** and Abhigyan Raj , *Barycentric defense*

- Venkatesh V, Md. Shad Akhtar, Mukesh Mohania, **Vikram Goyal**, *Auxiliary Task Guided Interactive Attention Model for Question Difficulty Prediction*

- Rishabh Gupta, Venktesh V, Mukesh Mohania, **Vikram Goyal**, *"'John ate 5 apples' != 'John ate some apples': Self-Supervised Paraphrase Quality Detection for Algebraic Word Problems"*

- Rajat Agarwal, Varun Khurana, Karish Grover, Mukesh Mohania, **Vikram Goyal**, *Multi-Relational Graph Transformer for Automatic Short Answer Grading*

- Astha Verma , **A V Subramanyam,** Rajiv Ratn Shah, *WASSERSTEIN METRIC ATTACK ON PERSON RE-IDENTIFICATION*

- Vineet Joshi, **A V Subramanyam**, *Contextual Active Learning for Person Re-Identification*

- S. Sharma and **A. Sharma**, *Multi-twisted additive codes over finite fields are asymptotically good, Cryptography and Communications.*

- Aditya Saini & **Ranjita Prasad**, *Select Wisely and Explain: Active Learning and Probabilistic Local Post-hoc Explainability*

- **Sarika Jain**, Pooja Harde, Nandana Mihindukulasooriya , *NyOn: A Multilingual Modular Legal Ontology for Representing Court Judgements*

- Aman Roy, Shravika Mittal, **Tanmoy Chakraborty**, *MG2Vec+: A Multi-headed Graph Attention Network for Multigraph Embedding, Knowledge and Information Systems*

- Subhabrata Dutta, Shravika Mittal, Dipankar Das, Soumen Chakrabarti, **Tanmoy Chakraborty**,: *Incomplete Gamma Integrals for Deep Cascade Prediction using Content, Network, and Exogenous Signals*

- Monika Yadav & **Anuradha Sharma**, *On the enumeration and classification of σ-LCD codes over finite commutative chain rings*

Year 2021

- Sandeep Sharman & **Anuradha Sharma**, *Multi-twisted Additive Codes with Complementary Duals over Finite Fields*

- Shivani Kumar, Anubhav Shrimal, Md Shad Akhtar, **Tanmoy Chakraborty**, *Discovering Emotion and Reasoning its Flip in Multi-Party Conversations using Masked Memory Network and Transformer*

- Hridoy Sankar Dutta, Kartik Aggarwal and **Tanmoy Chakraborty**, *DECIFE: Detecting Collusive Users Involved in Blackmarket Following Services on Twitter*

- Megha Sundriyal, Parantak Singh, Md. Shad Akhtar, Shubhashis Sengupta, **Tanmoy Chakraborty**, *DESYR: Definition and Syntactic Representation Based Claim Detection on the Web Link*

- Debajyoti Bera, Rameshwar Pratap, Bhisham Dev Verma, Biswadeep Sen, **Tanmoy Chakraborty**, *QUINT: Node embedding using network hashing*

- Shraman Pramanick, Shivam Sharma, Dimitar Dimitrov, Md. Shad Akhtar, Preslav Nakov**, Tanmoy Chakraborty**, *MOMENTA: A Multimodal Framework for Detecting Harmful Memes and Their Targets*

- Subhabrata Dutta, Tanya Gautam, Soumen Chakrabarti, **Tanmoy Chakraborty**, *Redesigning the Transformer Architecture with Insights from Multi-particle Dynamical Systems*

- Poorav Desai, **Tanmoy Chakraborty**, Md Shad Akhtar**, *Nice perfume, How long did you marinate in it?***

- Manjot Bedi, Tanisha Pandey, Sumit Bhatia, **Tanmoy Chakraborty**, *Why Did You Not Compare With That? Finding Papers for Use as Baseline*

- Priya Mehta, Meet Maheshwari, Brihi Joshi, **Tanmoy Chakraborty**, *BaDumTss: Multi-task Learning for Beatbox Transcription*

- Siddharth Sadhwani, Nishant Grover, Md Shad Akhtar, **Tanmoy Chakraborty**, *Detecting Anchors' Opinion in Hinghlish News Delivery*

- Subhabrata Dutta, Jeevesh Juneja, Dipankar Das, **Tanmoy Chakraborty**, *Can Unsupervised Knowledge Transfer from Social Discussions Help Argument Mining?*

- Shivani Kumar, Atharva Kulkarni, Md Shad Akhtar**, Tanmoy Chakraborty**, *When did you become so smart, oh wise one?! Sarcasm Explanation in Multi-modal Multi-party Dialogues*

- Shivani Kumar, Anubhav Shrimal, Md Shad Akhtar, *Tanmoy Chakraborty*, *Discovering Emotion and Reasoning its Flip in Multi-Party Conversations using Masked Memory Network and Transformer*

- Venktesh V, Mukesh Mohania and **Vikram Goyal** *, TagRec: Automated tagging of questions with hierarchical learning taxonomy*

- **Sarika Jain**, Sudipto Ghosh, Nandana Mihindukulasooriya, Abhinav Dubey, Pooja Harde, Ankush Bisht, *Constructing a Knowledge Graph from Indian Legal Domain*

- **Sarika Jain**, Deepak Jaglan, Kapil Gupta**, *Investigating the Similarity of Court Decisions***

- Siddharth Bhatia, Yiwei Wang, Bryan Hooi, **Tanmoy Chakraborty**, *GraphAnoGAN: Detecting Anomalous Snapshots from Attributed Graphs Link*

- Dhruv Sahnan, Snehil Dahiya, Vasu Goel, Anil Bandhakavi, **Tanmoy Chakraborty**, *Better Prevent than React: Deep Stratified Learning to Predict Hate Intensity of Twitter Reply Chains*

- Subhabrata Dutta, Samiya Caur, Soumen Chakrabarti, **Tanmoy Chakraborty**, *Semi-supervised Stance Detection of Tweets Via Distant Network Supervision*

- Ganeshan Malhotra, Abdul Waheed, Aseem Srivastava, Md Shad Akhtar, **Tanmoy Chakraborty**, *Speaker and Time-aware Joint Contextual Learning for Dialogue-act Classification in Counselling Conversations*

- Hridoy Dutta, Nirav Diwan, **Tanmoy Chakraborty**, *Weakening the Inner Strength: Spotting Core Collusive Users in YouTube Blackmarket Network*

- Hridoy Dutta, **Tanmoy Chakraborty**, *Blackmarket-driven Collusion on Online Media: A Survey*

- Venktesh V, Mukesh Mohania and **Vikram Goyal**, *Topic Aware Contextualized Embeddings for High Quality Phrase Extraction*

- Dhananjay Kumar Singha, Subrata Nandia, **Tanmoy Chakraborty,** Prasenjit Choudhurya**, *Disintegrating constant communities in complex networks***

Newsletter Editorial Team

Souravi Halder
Web Manager
iHub Anubhuti-IIITD
Foundation

Swati Mangla
Junior Manager
iHub Anubhuti-IIITD
Foundation