

CONTACT

☎ +91 8828381360
✉ tusharanand1594@gmail.com
● Bengaluru, India
📍 [tushar-anand1594](#)
© [tushar-anand15](#)
⊕ [tushar-anand.com](#)

CORE COMPETENCIES

- Generative AI
- Large Language Models
- Retrieval Augmented Generation
- AI Agents
- Machine Learning & Deep Learning
- NLP & Multimodal AI
- Vector Databases
- Information Retrieval
- Data Engineering & ETL Pipelines
- Backend & Distributed Systems
- Cloud Infrastructure
- Model Evaluation
- Prompt Engineering

FEATURED PROJECTS

- Legal RAG Platform — 15M+ docs
- Agent Memory System
- Gazette Knowledge Base
- AI Financial Validation Engine

SKILLS

Languages

Python, JavaScript/TypeScript, SQL

Backend & Infra

Django, Flask, FastAPI, REST, async, Celery, SQS, Docker, Kubernetes, Nginx, pytest, CI/CD, Git

Data & Storage

PostgreSQL, Neo4j, Redis, Elasticsearch, ETL pipelines, Selenium, BeautifulSoup

ML & NLP

PyTorch, TensorFlow, multilingual NLP, ASR/STT, VLMs, RAG, BM25, FAISS, agent evaluation, LangChain, Langfuse

Cloud

AWS EC2, S3, Lambda, GCP, Linux

SELECTED PUBLICATIONS

- [Inheritance Rights of Transgender Persons in India](#) — Indian Law Review, 2022
- [Problems with eCourts Data](#) — NIPFP Working Paper, 2020, on

Tushar Anand

Senior AI Engineer | LLMs, RAG & Agentic Systems

SUMMARY

Senior AI Engineer with 7+ years building production-scale LLM and machine-learning systems spanning agentic AI, retrieval-augmented generation, multimodal document processing, and large-scale data infrastructure. Have shipped retrieval systems over 15M+ documents, cut inference costs 75%, and automated expert workflows from weeks to hours. Blend deep backend and ML engineering with an applied-research background in quantitative methods.

CAREER HIGHLIGHTS

- Built AI systems processing 15M+ documents
- Reduced token costs by 75%
- Improved validation turnaround from one month to five hours
- Built multilingual AI for 8 Indian languages
- Winner — Agami Data for Justice Challenge 2019
- Publication in the Indian Law Review
- AI platform in demo at the Karnataka High Court

WORK EXPERIENCE

Senior Technical Consultant

June 2026 – Present

[Vidhi Centre for Legal Policy, Remote](#)

Applying AI and data engineering to make India's primary legal sources structured, searchable, and machine-readable.

- Designed and deployed a bilingual Tamil and English scraper and ingestion pipeline to normalize ~15,000 Tamil Nadu Gazette issues since 2008, enabling structured data access
- Developed a legislation tracker consolidating 970 acts with their full amendment histories into a machine-readable format, delivering a unified and accurate view of evolving laws for legal practitioners
- Built a RAG query layer over the normalized corpus using local embeddings and reciprocal rank fusion, exposed via an MCP server for structured querying by LLM agents

Founding AI Engineer

Jan – June 2026

[Superjoin.AI, Bangalore](#)

Founding engineer on the AI team, building the agentic harness and designing agent workflows across the platform.

- **DRHP Validation Platform — automated validation of Draft Red Herring Prospectuses, the preliminary IPO filings submitted to SEBI**
 - Engineered an end-to-end platform enabling automated validation of uploaded prospectuses and supporting documents
 - Automated “circle-up” by cross-verifying financial figures against CA certificates and restated statements, eliminating manual effort
 - Delivered analyst-grade first drafts at **88% accuracy**, reducing turnaround from one month to **five hours**
- **Superjoin Memory**
 - Designed a two-tier system: persistent user-level profiles and task-level memories recalled on demand by prompt and workflow
 - Exposed memory as a tool, routing retrieval through a relevance-judging model so only pertinent memories entered context
 - Reduced tool calls on long-running tasks by **10%**, improving agent efficiency
- **Evaluation**
 - Developed 15 handcrafted test cases spanning long- and short-running

data-quality issues across 1M+ cases

- [Gender Discrimination in Property Devolution under the Hindu Succession Act](#) — NIPFP Working Paper, 2020
- [The Unrealized Potential of Judicial Data in India](#) — Indian Express, 2020
- Advised the Delhi High Court eCourts committee headed by Justice Shakti Dharam on court modernization and computerization
- Co-authored reports evaluating the SWAMITVA Scheme for the Ministry of Rural Development and Right of Way in telecom with TRAI

CERTIFICATIONS

- TensorFlow Developer Professional, NLP, and Deep Learning Specializations — DeepLearning.AI
- Machine Learning Specialization — Stanford

EDUCATION

MA, Urban Policy and Governance

Tata Institute of Social Sciences, Mumbai
2017–2019

BA, Economics, Political Science and Sociology

Christ University, Bangalore
2014–2017

tasks to measure numeric and procedural accuracy

- Combined LLM-as-judge with precision and recall metrics, tracking every run to flag regressions instantly
- Cut new-model validation time from **three days of manual tuning to one day**, accelerating deployment cycles

AI Research Scientist

Oct 2024 – Dec 2025

Jhana.ai, Bangalore

Advanced AI research and product development across legal tech, courtroom automation, and multilingual transcription.

• Paralegal

- Led development of an [agentic legal chatbot](#) for research and drafting, improving speed and accuracy of legal workflows
- Built a hybrid retrieval system over **15M+ documents** using BM25, vector search, and knowledge graphs, enabling comprehensive query coverage
- Cut token costs by **75%** and quadrupled throughput on multimodal pipelines through VLMs and image processing
- Delivered a React/TypeScript frontend with click-through citations, smart highlights, and context navigation

• Courtroom Project

- Directed end-to-end development of an AI-enabled [courtroom automation platform](#), now in demo at the **Karnataka High Court**
- Built ingestion pipelines for **2,000+ daily case dockets**, streamlining e-filing flows
- Automated headnote generation — concise summaries of a judgment's key legal holdings that let lawyers gauge relevance without reading the full order; in a study with 200 lawyers and law students, outputs were preferred **80% of the time**

• National Judicial Data Pipeline

- Designed and operated an ingestion backend scraping all judicial sources at **15M+ document scale**
- Automated daily refresh of Elasticsearch and FAISS indices, ensuring up-to-date access for **5,000+ users**

• Steno

- Spearheaded development of a [transcription tool](#) tailored for Indian lawyers
- Tuned STT APIs and open-source ASR models for **8 Indian languages** and legal vocabulary
- Optimized performance for noisy courtroom recordings and variable speakers, improving transcription reliability

Research Consultant — Legal Systems Team

Jan 2022 – Oct 2025

XKDR Forum, Mumbai

Applied data engineering and statistical methods to analyze judicial processes, improving transparency and efficiency in case-flow research.

- Built an [ETL pipeline](#) to extract and structure commercial case data from the Bombay High Court, enabling systematic analysis
- Scraped and processed **5,000+ PDF orders** from the Bombay High Court and Debt Recovery Tribunals, creating a usable dataset for research
- Applied deep learning models to classify case orders as substantive or non-substantive, achieving **85% accuracy** and reducing manual review
- Conducted [survival analysis](#) to model case-disposal and hearing timelines, generating insights into judicial delays and efficiency

Research Fellow

Mar 2019 – Nov 2021

National Institute of Public Finance and Policy, New Delhi

Led legal data projects combining NLP, text mining, and econometric analysis to support justice reform and fiscal policy research.

- Won the [Agami Data for Justice Challenge 2019](#) by building and openly releasing a **1M-case dataset** from district courts, hosted at [The Justice Hub](#); findings presented to the Delhi High Court e-Courts Committee and XKDR Legal Systems seminar
- Extracted and processed **6,731 ITAT orders** for a [transfer-pricing study](#) using NLP and regex-based text mining, enabling structured analysis of tax jurisprudence
- Analyzed **NASA VIIRS nightlights and forest-fire satellite data** as remote-sensing proxies for economic activity, integrating geospatial signals into fiscal and development policy research