

IndraAstra.AI

Power of Enterprise grade AI at the cost of a Database

About Me : Dhiraj Daga (Half Scientist - Half Engineer)

(<https://www.linkedin.com/in/dddhiraj>)

Notable work

- Leading AI at Allen Digital one of the largest education company in India (300K+ active users)
- Boosted Walmart's fresh produce defect detection accuracy by 8% through automated model architecture optimization
- Solved document structure recognition and OCR for Project Anuvaad, a document to document translation tool currently used by the Supreme Court of India to fast-track Judgement translations. (<https://www.anuvaad.org>)
- Created India's largest Indic OCR validation dataset (<https://arxiv.org/pdf/2205.02543>)
- Facilitated creation of largest set of indic parallel sentences (<https://arxiv.org/pdf/2104.05596>)



Immediate Value - AI Performance Engineering

What We Deliver

Custom optimization for businesses already using AI (we want to steer away from laggards who have not started using AI yet)

Three Performance Levers (Based on Your Business Priority):

1. **Cost Reduction** - Lower operational expenses of using AI for better margins or competitive pricing
2. **Latency Optimization** - Faster inference for user-facing applications where response time matters
3. **Accuracy Enhancement** - Improved reliability for high-stakes use cases where errors are costly

Engagement Model

- **2-month POC** - Fast validation, low commitment
- **Expansion** - Scale to full implementation based on POC results
- **End-to-end ownership** - From business analysis to deployment

Long-Term Vision - Next-Generation LLM Architecture

Novel Algorithm Under Development (Model to be Demo ready in early half of 2026)

1. **1000x Energy Efficiency** - reduce the cost of hardware and electricity for data centers
 - Theoretical validation complete
 - Fundamentally different information processing paradigm
2. **Database-Like Scaling Economics** - LLMs whose cost economics mimic the cost brackets of scaling databases
 - Hyper distributed compute which can leverage clients hardware to fulfil client's inference need, so that there is no centralised bottleneck as number of users increase
3. **Independent LLM Provider** - Building toward Anthropic-equivalent domain specific information processing models.