

# TAYYAB M.

📍 Nancy, France 📞 +33 7 52 92 75 42 📩 hello@tayyab.io  
🌐 tayyab.io 💬 tayyab.io/in 🐾 tayyab.io/github

## Profile

Final-year Master's (M2) student in NLP at Université de Lorraine building practical, reliable, and measurable LLM-powered systems. Experience in **LLM fine-tuning, RAG pipelines, HITL annotation workflows, and model interpretability** across the full ML lifecycle. Seeking a **6-month End-of-Studies Internship (stage de fin d'études)** in Applied Science / NLP Engineering. Available early March 2026.

## Education

<b>Université de Lorraine (IDMC)</b> <i>MSc in Natural Language Processing (M2)</i>	<b>Nancy, France</b> <i>2024 – 2026</i>
– Deep Learning for NLP, Generative AI (LLMs), Prompt Engineering, Information Retrieval	
<b>University of the People</b> <i>BSc in Computer Science – CGPA: 3.80/4.00</i>	<b>Pasadena, CA, USA</b> <i>2020 – 2023</i>
– Focus: Data Science, Machine Learning, Software Engineering	

## Experience

<b>Researchlytic</b> <i>Creator &amp; Full Stack Engineer</i>	<b>Remote</b> <i>2022 – 2024</i>
– <a href="#">Live Demo</a>   <a href="#">Video Demo</a>   <a href="#">Source Code</a> – <b>Built a bibliometric analytics platform on top of OpenAlex</b> (250M+ works, 90M+ authors, 100K+ institutions) to explore research trends across institutions, countries, and time. – <b>Achieved fast UX and analytics performance:</b> Lighthouse score <b>96</b> , <b>0.7s FCP</b> , <b>0.8s LCP</b> ; dashboard queries compute metrics and render results in under <b>3 seconds</b> . – Developed ranking pipelines combining <b>publication volume, citation impact (SCR/SCRP)</b> , and <b>open-access signals</b> to identify emerging high-impact research domains.	

## Projects

<b>AnnotaLoop</b> — LLM-Driven Document Annotation with HITL   <i>M2 Software Project</i>
– <a href="#">Project Website</a>   <a href="#">Video Demo</a>   <a href="#">Source Code</a>
– Reduced manual annotation time from <b>approx. 30 min to approx. 5 min per document (83% reduction)</b> using <b>under 1 minute</b> LLM-generated suggestions with human-in-the-loop review, combining LLM efficiency with human validation to produce <b>structured labeled data at scale</b> .
– Built a <b>cross-platform desktop app</b> (macOS, Windows, Linux) integrating <b>Mistral</b> , <b>OpenAI</b> , <b>Anthropic</b> , <b>Gemini</b> , <b>Ollama</b> , and <b>LM Studio</b> with batch processing, offline-first privacy, and exports to <b>JSON</b> , <b>CSV</b> , and <b>annotated PDFs</b> .

### Intrinsic Evaluation of French Word Embeddings | *M2 Supervised Research*

– <a href="#">Live Demo</a>   <a href="#">Video Demo</a>   <a href="#">Source Code</a>
– Analyzed grammatical gender encoding across <b>33K+ French nouns and adjectives</b> using FlauBERT embeddings and attribution methods ( <b>SHAP</b> , <b>LIME</b> , <b>RF</b> , <b>EKA</b> ).
– <b>Demonstrated distributed (non-localized) encoding:</b> achieved <b>49–85% accuracy with only 1% of dimensions</b> and enabled up to <b>75% dimensionality reduction with &lt;5% accuracy loss</b> .
– <b>Ran 16 configurations</b> (4 models × 2 classifiers × 2 tasks) with peak <b>99% adjective accuracy</b> and strong within-model convergence in <b>15/16 configurations at 5% threshold</b> .

 [View portfolio:](#) Project overviews, live demos and video demos at [tayyab.io](#)

## Technical Skills

**GenAI & LLMs:** LLM Fine-tuning, RAG Pipelines, LangChain, Prompt Engineering, BERT, LLaMA, GPT APIs  
**ML/NLP:** PyTorch, Hugging Face, scikit-learn, spaCy, SHAP, LIME  
**Engineering:** Python, SQL, PHP, JavaScript, Docker, Git/GitHub, AWS (S3), Linux  
**Vector DBs:** FAISS, Pinecone, Chroma  
**Languages:** English (Fluent), Urdu/Punjabi (Native), French (Basic)