

Hicham Yassin

Multilingual NLP & Model Evaluation Engineer

Nantes, France & Rabat, Morocco | Mathis.yassin@gmail.com | +33 7 63 26 26 67 | github.com/h-yassin-ai

NLP

OCR/HTR

RAG

CORE SKILLS

- Multilingual NLP
- Model evaluation
- Prompt design
- Error analysis
- Linguistic QA

DOCUMENT AI

- OCR, HTR, HMER
- Arabic manuscripts
- Dataset curation
- OCR-to-translation

ENGINEERING

- Python, PyTorch
- Hugging Face
- Transformers, VLMs
- LoRA / QLoRA
- FastAPI, Docker
- RAG chatbots
- Linux, Git
- Next.js, React, TS

LANGUAGES

- French - native
- English - C1
- Literary Arabic - advanced

PUBLISHING

- Translation
- Writing
- Typesetting
- Book covers

INTERESTS

- Chess & logic
- Arabic poetry
- Typography
- Philosophy of science

LINKS

yassinhicham.omnignosis.fr
github.com/h-yassin-ai
khattgpt.vercel.app

ROLE FIT

Adaptable engineer
Research + shipped tools

PROFILE

Multilingual NLP engineer working across French, English and Arabic. Builds evaluation workflows, OCR/HTR pipelines, datasets, RAG chatbots and web tools, with a research background in NLP and document AI.

EXPERIENCE

Scale AI - AI Training & Model Evaluation Contributor

Remote | May-Aug 2025

- Reviewed prompts and model outputs for AI training/evaluation.
- Checked reasoning, instruction-following, response structure and output quality.
- Applied QA rubrics in iterative review workflows.

Freelance Web Developer & RAG Chatbot Integrator

Independent projects

- Built websites, internal tools and AI assistants with React/Next.js, Python/FastAPI and API integrations.
- Integrated RAG chatbots for document search, retrieval workflows and domain-specific Q&A.
- Delivered end-to-end projects from client need to deployment, content structure and maintenance.

Editions Heritage Mohammadien - Founder, Author, Translator & Publisher

2019-Present

- Translate, write and publish Arabic-French books through an independent publishing house.
- Handle book typesetting, cover design, web publishing and bilingual editorial production.
- Use AI-assisted OCR-to-translation workflows for publishing and bilingual dataset creation.

SELECTED PROJECTS

Khatt-GPT - Collaborative HTR Platform

Test: khattgpt.vercel.app

- Web platform for manuscript transcription, annotation and HTR dataset curation.

URL: <https://khattgpt.vercel.app/>

Structure-Aware Historical HMER - Leibniz Manuscripts

PHILIUMM / LS2N

- Adapted CoMER and Uni-MuMER-Qwen3.5-2B with QLoRA and structural supervision.
- Reached 73.05% ExpRate and 0.95 token edit distance on the 961-expression S-C test split.
- Manuscript in progress; error analysis on omissions, scope, decorators and rare notation.

Code: gitlab.univ-nantes.fr/E197486J/leibnitz

Demo: github.com/h-yassin-ai/Leibniz-HTR-HMER

Historical Maghrebi HTR - Private Project

Arabic manuscripts

- Fine-tuned Qwen 3.5 9B on ~11K Arabic manuscript lines; built OCR/HTR data pipelines.

Demo: <http://htr.omnignosis.fr/manuscript>

AjurrumAI - Arabic Grammar Tutor

SDAIA Hackathon

- Fine-tuned ALLAM for Arabic syntax and translation; built a grammar UI with syntax trees.

Arabic Companion - Conversational Language Tutor

Llama 3 Hackathon

- Built an Arabic learning assistant with Llama 3, LlamaIndex, Together AI and LangChain.

Arabic-OCR - Work in Progress

Current project

- Building scanned-manuscript preprocessing, OCR-to-translation experiments and dataset tooling.

EDUCATION

MSc Computer Science - NLP & Smart Computing (ATAL)

Nantes Universite

Double BSc Major in Mathematics & Economics

Universite du Mans

KEY STRENGTHS

Evaluation, document AI and web delivery across multilingual, research-heavy projects.

Model Eval

multilingual evaluation

Doc AI

OCR / HTR / datasets

Builder

research to product