

Ustunpasa Igdir

LLM Evaluator | AI Data Annotation Specialist | STEM & Multilingual QA Reviewer

Barcelona, Spain | contact@ustunpasaigdir.com | www.ustunpasaigdir.com | LinkedIn: [linkedin.com/in/ustunpasaigdir](https://www.linkedin.com/in/ustunpasaigdir)

Profile

Physics graduate with 7 years of combined experience in translation, linguistic QA, AI data annotation, and freelance evaluation projects, including work with Datamundi, Blend, DataAnnotation, and other project-based clients. I evaluate LLM and AI model outputs for correctness, instruction following, hallucinations, factual errors, incomplete answers, and reasoning quality.

My current focus is practical model-output QA for STEM / physics tasks, multilingual AI data review, and rubric-based annotation. Turkish is my native language; I work with advanced professional proficiency in English and working proficiency in Spanish. I am comfortable applying detailed guidelines, handling ambiguous outputs, and writing short reviewer notes for edge cases.

Core Services

- AI model response evaluation
- STEM and physics answer review
- Turkish-English-Spanish AI data QA
- Rubric-based annotation and labeling
- Gold-answer comparison
- Error and hallucination tagging
- Dataset review and cleanup

Relevant Experience

AI Data Annotation, Linguistic QA & Freelance Evaluation Work - Remote

Datamundi, Blend, DataAnnotation, and project-based clients | 2019-Present

- Reviewed and annotated multilingual data according to detailed project guidelines.
- Worked on translation, linguistic QA, AI data annotation, model-output review, and quality evaluation projects.
- Checked accuracy, relevance, fluency, completeness, consistency, instruction following, and reasoning quality.
- Flagged vague reasoning, missing information, factual issues, unnatural wording, hallucinations, and guideline violations.
- Applied Turkish-English-Spanish language judgment in multilingual review and AI data QA tasks.

Freelance / Project-Based LLM Evaluation Work - 2025-Present

- Reviewed LLM outputs for correctness, instruction following, reasoning quality, and hallucinations.
- Evaluated STEM and physics-style answers for numeric accuracy, assumptions, and conceptual clarity.
- Created and applied practical labels for model-output QA, dataset review, and human review.

Portfolio Case Study

LLM Scientific Reasoning Evaluation

A portfolio-scale case study evaluating whether language model answers preserve correct math, assumptions, and reasoning structure in scientific tasks.

Demonstrates:

- Numeric and conceptual error detection
- Assumption-handling review
- Failure-mode labeling
- Human review after automatic checks
- Follow-up robustness testing

Portfolio: <https://ustunpasaigdir.com>

GitHub: <https://github.com/ustunpasaigdir/llm-scientific-reasoning-eval>

Education

BSc Physics, Istanbul University, 2018

GPA: 3.49 / 4.00

Undergraduate project implementing a quantum circuit on IBM Quantum hardware.

Languages

- Turkish: Native
- English: Advanced professional proficiency
- Spanish: Working proficiency