

# Santiago Tomas Torres

Email: santiagotomastorres7@gmail.com

Bachelor of Engineering

University of San Andrés

Phone: +542216360514

[GitHub Profile](#)

[LinkedIn Profile](#)

I am a final-year student with a solid technical background in Python, C/C++, and emerging AI technologies. I developed strong soft skills in communication, teamwork, and adaptability through research and collaboration in interdisciplinary and inter-university teams, participating in demanding, high-impact projects. Furthermore, I adapt quickly and easily to different topics and work environments.

## EDUCATION

---

- **B.Eng. in Artificial Intelligence Engineering (GPA: 8.14/10)** *Mar 2022 – Present*  
*University of San Andrés* San Fernando, Buenos Aires, Argentina
  - Student of the first cohort of the country's first Artificial Intelligence Engineering program. Currently in the final year of the degree.

## EXPERIENCE

---

- **Research Team Member** *Mar 2025 – Dec 2025*  
*University of San Andrés* San Fernando, Argentina
  - Photorealistic synthetic dataset generation with automated capture and labeling for future YOLO training; combined Nvidia Isaac Sim and ROS2 with simulated drone navigation.
  - Tools: Nvidia Isaac Sim, ROS2.
- **Research Team Member** *Nov 2024 – Jan 2026*  
*University of San Andrés – Northwestern University* San Fernando, Argentina
  - Co-authored and published the resulting research paper: *Linnaeus: A Hierarchical, Multi-Label Framework for Autonomous System Classification* (arXiv:2603.13649).
  - Joint project with Northwestern University on taxonomic classification of Autonomous Systems (AS) on the Internet.
  - Fine-tuning of LLMs for multiclass classification.
  - Tools: LLMs, OpenAI API fine-tuning jobs, Web scraping, Data Labeling.
- **Research Team Member** *Feb 2024 – Dec 2024*  
*Interdisciplinary Laboratory of Time and Experience (LITERA) – UdeSA* San Fernando, Argentina
  - Prediction of sleep event start/end at individual and population levels using Android activity data (Google Takeout); wearable actigraphy used as ground truth for validation.
  - Tools: Machine Learning, Deep Learning, Data Collection.

## PUBLICATIONS

---

- **Linnaeus: A Hierarchical, Multi-Label Framework for Autonomous System Classification** *Mar 2026*  
*arXiv preprint [cs.NI]*
  - Piotto, M.; Schuemer, I.; Torres, S. T.; Beiró, M. G.; Carisimo, E.; Bustamante, F. E. [arXiv:2603.13649]

## PROJECTS

---

- **Text-to-SQL Assistant with Local LLMs** *Jan 2026*  
*GenAI Text-to-SQL system for sales* Personal project
  - Designed and implemented an end-to-end Text-to-SQL pipeline with local LLMs (Ollama), converting natural language queries into executable SQL and narrative answers over sales data in PostgreSQL.
  - Developed structured prompts with schema injection, business rules, and schema linking; implemented a two-stage inference flow (SQL generation + natural language summary) and a self-correction loop.
  - Added security guardrails (SQL sanitization, read-only validation) and result grounding to reduce hallucinations in small models.
  - Orchestrated the solution with FastAPI, SQLAlchemy, and PostgreSQL; containerized the architecture with Docker Compose (Streamlit frontend, backend, database, model server) for reproducible deployments.
  - *Technologies: Python, FastAPI, Streamlit, PostgreSQL, Ollama, Docker, Docker Compose, Prompt Engineering, Text-to-SQL, LLM Evaluation.*

- **Soccer Commentator**

Jun 2025

Real-time soccer commentary generation with NLP and distributed microservices.

University of San Andrés

- Implemented an LLM/NLP workflow with text-to-speech and distributed components.
- Fine-tuned an open-source text-to-speech model using LoRA adapters (PEFT library) for voice customization.
- Integrated components through gRPC and FastAPI, with deployment in Docker containers.
- *Technologies: OpenAI API, Azure OpenAI, Ollama, FastAPI, Docker, Whisper, gRPC, ElevenLabs.*

- **Wildlife Monitoring with Thermal Images**

Dec 2024

Automated detection of wild animals using RGB and thermal drone imagery.

University of San Andrés

- Trained YOLOv5/YOLOv11 models for multi-modal detection.
- *Technologies: Computer Vision, Machine Learning, YOLO, Deep Learning.*

- **Prediction of MercadoLibre SUV Prices**

Jun 2024

Estimation of SUV prices from Mercado Libre listings.

University of San Andrés

- Built supervised models to predict prices; explored gradient boosting and deep learning baselines.
- *Technologies: Machine Learning, XGBoost, Deep Learning.*

## TECHNICAL SKILLS – STACK

---

**Programming Languages:** Python; C/C++; Java

**Python Libraries:** NumPy, Pandas, Matplotlib, PyTorch, scikit-learn, TensorFlow, OpenCV, XGBoost

**Databases and Storage:** SQL (PostgreSQL), MongoDB, Apache Spark

**LLMs and NLP:** OpenAI API, Azure OpenAI, NLTK, spaCy, RAG

**Frameworks and Deployment:** FastAPI, Streamlit, Docker, Docker Compose

**Robotics:** ROS2

**Tools:** GitHub Copilot, Cursor, Claude Code, Codex, Gemini CLI, VS Code, IntelliJ IDEA, L<sup>A</sup>T<sub>E</sub>X

**Operating Systems:** Linux, Windows

## LANGUAGES

---

- **Spanish** (Native)
- **English** (B2 / C1)

## ACHIEVEMENTS

---

- **1st Place Hackathon** Siemens & Louis Dreyfus Company & Union Industrial Argentina *Oct 2025*
  - Optimization in the soybean grain sampling process
  - Pitch focused on Computer Vision and Machine Learning.
- **1st Place HackITBA** Sustainability Category *Mar 2025*
  - Platform that uses NLP to analyze and compare unstructured ESG reports from companies, facilitating the assessment of their environmental, social and governance impact, allowing direct comparison with other companies in terms of sustainability.
  - Skills: Communication, NLP, and LLMs solution stream.
- **2nd Place AI E-Commerce Challenge** (Amazon Web Services (AWS), Trafilea, Mutt Data) *Mar 2024*
  - Platform that uses Generative AI with LLMs for personalized ads using user data.