

Jose Antonio Tolentino Veliz

joseatolentinov@gmail.com | joseatolentino.com | github.com/jatolentino | linkedin@joseatolentino | Lima, Peru | +51 997008163

EXPERIENCE

Scryper, New Mexico, USA Jan. 2022 - Present

Software Engineer (remote)

- Integrate Large Language Models (LLMs) into web and mobiles applications, facilitating advanced search and recommendation features that led to a 30% increase in user engagement (OpenAI, Spring Boot, Django, and Flutter)
- Develop interactive and intuitive user interface using React and Redux for a data visualization platform and implement dynamic components, and user-friendly navigation, resulting in a 25% decrease in bounce rates
- Leverage NLP techniques to develop intelligent chatbots for customer support, reducing response times by 50% and improving customer flow by 25% (OpenAI API)

QbAI, Lima, Peru Jan. 2020 - Dec. 2021

Full Stack Engineer

- Led the development of a robust RESTful API using Django and Python for a SaaS platform, achieving an average response time of under 100ms and handling over 10,000 requests per minute. Implemented caching mechanisms and load balancing strategies to ensure high availability and fault tolerance of the system
- Implemented a scalable backend system using Java and Spring Boot for an e-commerce platform, resulting in a 50% increase in concurrent users and a 30% decrease in response time. Employed database indexing and caching techniques to optimize query performance, enhancing overall system efficiency

Automation Lab, Lima, Peru Jan. 2017 - Dec. 2019

Control and Backend Engineer

- Integrated FPGA boards with Kubernetes clusters to enable high-speed data processing for real-time control systems and leveraged FPGA's hardware-level parallelism and Kubernetes' orchestration capabilities to achieve sub-millisecond response times, improving system responsiveness and ensuring precise control in time-critical applications
- Designed and implemented a scalable edge computing infrastructure using Kubernetes and FPGA-based edge devices, resulting in a 40% reduction in data transfer latency and improved real-time decision-making capabilities
- Implemented a highly available and fault-tolerant IoT platform using Kubernetes and Docker, resulting in a 50% decrease in system downtime and improved reliability of data collection from thousands of IoT devices

EDUCATION

New York University Sep. 2022

MSc. in Computer Engineering (Dropped out)

National University of Engineering Mar. 2012 - Jul. 2018

BS. in Mechatronics Engineering

Thesis: "Design, and Implementation of a Control Strategy for Non-linear Models Applied to an Industrial Process in a Multi-Tank System, Based on a Genetic Algorithm"

RESEARCH EXPERIENCE

Information Technology, and Communications Center (CTIC-UNI), Lima, Peru Jan. 2016 - Aug. 2016

Embedded Systems Intern

- Developed and prototyped printed circuit boards for robotics applications

CanSat Leadership Training Program, University of Hokkaido, Japan Apr. 2016 - Oct. 2016

- Developed, launched and deployed a nano satellite in a can that performs telemetry at 500ft

ARLISS Competition, Nevada, United States Sep 2016 - Oct. 2016

- Developed a miniaturized rover that navigates autonomously to a GPS predefined target in the desert

## PROJECTS

---

**Rover:** Autonomous rover that can avoid obstacles and navigate to a GPS coordinate. Status: completed  
**CanSat:** Nano satellite in a can that performs telemetry. Status: completed  
**LittleVid:** Video to Image and Text summarization. Status: beta  
**DBot:** AI chatbot that can query product's stock on the database and reply back to users. Status: completed  
**Rail:** Electric Autonomous Urban VTOL for a single passenger with mobile app. Status: on development  
**SmileMe:** Transfer a smile expression to a non smiley face photo. Status: beta  
**Faang150:** Draw and code on a web IDE with a stylus pen on touch screen devices. Status: alpha  
**AutoZoom:** Apply automatic zoom on desired areas in an X-ray surgery procedure. Status: on development  
**NERF:** Social Network that uses NERF technologies to animate face photos. Status: on development  
**DubbAI:** Dubbing videos transferring emotion from one language to another with lip sync. Status: on development  
**UnrealCode:** Generate front-end and back-end code components from graphics design. Status: on development

## PATENT

---

**IoT Industrial Liquid Level Control System, 002191-2019/DIN** **Issued: April 2021**  
 Apparatus that controls liquid level in a cascade tank system by means of an FPGA and enables users to implement their own control strategy or AI algorithm approach. Valued at 100 thousand USD.

## PAPERS

---

**International Journal of Mechanical and Production Engineering, ICRAMM** **Published: July 2021**  
 Title: LPV-MPC control, and self-tuning feedback gains for the trajectory tracking of a quadcopter UAV

**2022 IEEE/SICE International Symposium on System Integration** **Published: February 2022**  
 Title: Control design for a multi-rotor VTOL enhanced by a Gradient Descent algorithm to optimize the position tracking

## CERTIFICATIONS

---

### Coursera, Google Cloud

Introduction to Generative AI	Introduction to Generative AI Studio	Introduction to Image Generation
Introduction to Large Language Models	Transformer Models and BERT Model	Create Image Captioning Models

### Coursera, DeepLearning.AI

Machine Learning Specialization	TensorFlow Developer Professional Certificate
TensorFlow: Data and Deployment Specialization	TensorFlow: Advanced Techniques Specialization
Machine Learning Engineering for Production	Deep Learning Specialization

### Red Hat

Red Hat Certified Engineer (RHCE)  
 Red Hat Certified System Administrator (RHCSA)

### Axelos Scrum

IT Infrastructure Library - ITIL V4 (ID: GR671224267JT)

## OTHERS

---

**Languages:** English (professional), Spanish (bilingual), and French (intermediate)