

Vinícius de Oliveira Civali

Software Engineer

Site viniciuscivali.com
E-mail viniciuscivali@gmail.com

Professional objectives

As a software engineer with strong experience in backend development, distributed systems, and automation, I'm particularly interested in the intersection of software engineering, backend, and AI. I'm eager to contribute by applying my experience with large-scale systems, code evaluation, and AI agent design to improve LLM performance on real-world engineering tasks. My background includes code quality enforcement, technical mentorship, and designing scalable tools — all of which align with the curation and evaluation of code datasets for AI models.

Professional experience

2025 - Current Position

Senior Software Engineer: UpWork

Experiences:

- **AI Recommendation System:** Orchestrating recommendation pipelines to connect talent with relevant projects, enhancing the platform's matchmaking experience.
- **Sourcing Expansion:** Successfully delivered the expansion of the Sourcing system, enabling AI agents to identify qualified freelancers for job posts from entry-level clients at scale.
- **Backend Development:** Building and maintaining scalable services using Java and Python, with a focus on performance and AI model integration.
- **Full-stack Contribution:** Implementing frontend adjustments (JavaScript) to ensure a seamless experience between recommendation logic and user interface.

Technologies:

- Java

- Python
- JavaScript
- AI/ML Integration
- Distributed Systems
- Backend Architecture.

2020 – 2025

Software Engineer: Amazon

Experiences:

- AI and AI Agents:
 - Design of **AI agents** to facilitate investigations into tax inconsistencies for our clients and leaders.
 - Design of **AI agents** to facilitate the team's passage of context, both business rules and infrastructure.
 - The first team focused on AI agent solutions for productivity within Amazon ITECH LATAM. I am one of the key contributors helping the team contextualize itself for this next step.
- Automation of tasks with Python scripts that impact the work of others in the team:
 - Automatic pagination of queries in the database, with no search limit per page.
 - Creation of images to visualize transaction graphs, to facilitate technical and product discussions, meetings, and designs.
- Implementation and mapping of the team's technical debts.
- Monitoring and supporting interns and entry-level.
- Refactoring old systems.
- System design.
- Coordinating the expansion of the team's product (Transaction data orchestration portal for tax audits) to European countries.
- Managed critical tasks during the country's tax reform and change of completeness rules, while the production stage was running.
- Developed and implemented filters to exclude invalid data in third-party seller operations.
- Executed high-risk Production stage procedures (MCMs).
- Created technical documentation and automated processes for report generation and acceptance testing.
- Activated and configured alarms to monitor operations during market expansion in other countries.
- Implemented improvements in scripts and processes, optimizing data visualization and analysis.
- Served as an on-call engineer during recurring periods, autonomously and effectively addressing issues outside standard working hours.
- Strong collaboration skills in international settings

Technologies:

- Java, TypeScript, Python, Kotlin;
- AWS (S3, DynamoDb, Lambda, Pipelines, CDK, SDK, EC2, ECS, ...);

- Ai Agent
 - Lang chain
 - Docling
 - Bedrock
 - OpenSearch
 - RAG
 - Embedding
 - Prompt engineering
 - Automation/Productivity/IA solutions
- Documentations;
- Object-oriented
- Domain-driven design
- System Design;
- Code Review;
- Elasticsearch;
- Back-end;
- Improving code organization;
- Performance concerns;
- Extensible code;
- Git;
- On-call experience;
 - Keep the system functional even outside of conventional working hours, dealing with complex problems independently.
- Microservice
- Distributed systems
- Mentoring entry-levels
- Mentoring interns
- DevOPS:
 - Pipelines CI/CD
 - Unit test
 - Integration tests
 - Load tests
- Scalability

2025 – 2025

Cloud AI Software Engineer consultant: [Cloud303.io](https://cloud303.io)

Experiences:

- **WAFR Automation AI review project:** Development of an automated, well-architected review product that analyzes the customer's AWS account and identifies all good practices and opportunities. EC2 running Docker with LangGraph connected to Bedrock and RAG using Knowledge Bases + OpenSearch.
- **TimescaleDB implementation + migration:** Classic blue/green end-to-end implementation of an EC2 cluster with Docker running TimescaleDB with vertical write scalability and horizontal read scalability.

- **LangGraph Pipeline for Generating Patient Health Reports:** Development of a FastAPI service with a load balancer, using DynamoDB and caching frequent requests. API connecting the LangGraph pipeline with the patient history RAG (using KnowledgeBases + OpenSearch) and with S3 storing the report model, resulting in a document with the patient history and next steps in treatment.

Technologies:

- LangChain.
- FastAPI.
- Python
- Docker.
- EC2.
- NLB.
- SG.
- ASG + NLB.
- VPC/Subnets.
- Bash Script.
- Bedrock.
 - Titan model
 - Titan Embedding
- Python.
- E2E Tests.
- Unit Tests.
- Embedding with Knowledge bases
 - Open Search
 - S3
 - Titan Embedding

Academic background

Fev 2017 – Jun 2024

UNICAMP — Campinas State University — Brazil
Engineering of Control and Automation.

Fev 2017 – Dec 2024

UNICAMP — Campinas State University — Brazil
Sound Engineering focused on Machine Learning.

Jun 2014 – Dec 2015

Technical School Duque de Caxias
Mechatronics technician.

Fev 2013 – Dec 2015

ETECAP – State Technical College Centro Paula Souza
High school.

Extracurricular activities / Personal projects

Fev 2025 – Marc 2025

AI agent attendant (freelancer)

A store selling imported products had problems updating the status of orders for its customers. As it was a small store, human service was insufficient, since the products came from another country, and it took a long time to arrive. Consequently, customers asked numerous questions.

I created an integration between the **WhatsApp API**, the **OpenAI API** (I used an intermediate model to reduce the cost, since this task wasn't so complex and didn't require 4.1.), and the **Shopify API** to implement an **AI agent** to solve the problem.

I used **Python** with **Poetry** to orchestrate dependencies, **the LangChain library** to create a chain of thought for the agent more easily, and also to make it easier to implement new **Tools** in the long term and make the application **scalable**.

As a **result**, the **demand for human agents fell dramatically**, and the **solution was able to help even customers who weren't familiar with technology**, as I implemented an agent with a more human prompt.

Experiences:

- Lang Chain
- Python
- Poetry
- WhatsApp Api
- OpenAI Api

Jan 2023 – Dec 2023

Academic Project: Real-time Guitar Music Transcription *University of Campinas (UNICAMP) - Machine Learning (ML)*

Developed a real-time music transcription system for acoustic guitar, focusing on transforming audio signals into symbolic music representations (sheet music/tablature) using **machine learning** and **signal processing** techniques.

Experiences:

- **Machine Learning & Neural Networks:** Implemented advanced machine learning models, including **LSTM** and **BiLSTM** networks, to enhance transcription accuracy and effectively process temporal data for real-time music transcription.
- **Signal Processing:** Applied techniques such as Short-Time Fourier Transform (STFT) and Mel-spectrograms to analyze and extract audio features, enabling precise real-time transcription of acoustic guitar notes.
- **Custom Dataset Creation:** Created a unique dataset of acoustic guitar MIDI recordings by curating and compiling data from various online sources, addressing the absence of existing datasets for this specific use case, **using Python to prepare the database**.

- **Data Handling & Model Training:** Pre-processed audio data and trained machine learning models on a diverse set of guitar recordings, ensuring model robustness across various musical styles and techniques.
- **Research & Innovation:** Conducted research on unsupervised music source separation and explored innovative techniques to improve transcription quality using differentiable parametric models.
- **End-to-End Development:** Designed, implemented, tested, and evaluated the entire system, from data ingestion and model training to MIDI transcription and visualization, providing an automated tool for real-time music transcription.

Technologies:

- Python (Keras, PyTorch), JavaScript (for visualization tools)
- LibROSA (audio analysis), PyAudio (real-time audio input), RtMidi (MIDI handling)
- LSTM, BiLSTM, Neural Networks, STFT
- Custom dataset creation, pre-processing, and feature extraction
- Full system design and architecture for real-time processing and transcription
- Transcription output via MIDI, interfacing with Logic Pro X for real-time music score visualization
- Signal processing, music source separation, and unsupervised learning techniques

Jan 2021 – Feb 2021

Global Game JAM 2021 development marathon participant

Game development marathon, performed by a team of 5 people; I participated as a Game developer and Game designer;

Technologies:

- C#;
- Back-end to Front-end relationship;
- Code organization;
- Work as a team in a short and stressful period.

Set 2020 – Set 2021

Undergraduate research in voluntary modality: PIBIC

(Institutional Program of Research Fellowships — UNICAMP) financed by the bank **Santander** through “Santander Entrepreneurship Training Program”.

Department of Systems and Energy DSE of the Faculty of Electrical and Computer Engineering (FEEC / UNICAMP).

Research title: Algorithm for forecasting the location of buses on the circular lines of Unicamp's campus through AI. - Machine Learning (ML)

Advised by Dr. Madson Cortes de Almeida

Experiences:

- **Machine Learning (ML):** Developed an **ML-based algorithm using artificial neural networks** to predict the location of buses on Unicamp's campus in case of communication failures.
- **Processed historical GPS data and real-time signals using Python** to estimate future bus positions, ensuring reliable location data even during signal outages.
- Improved the user experience and transport management by providing more accurate and continuous bus tracking information.

Technologies:

- Artificial intelligence;
- Python;
- Data processing;
- Big data;

Apr 2020 – Apr 2020

April game development marathon (UNICAMP's game development extracurricular - GAMUX)

Game development marathon, performed by a team of 5 people. I participated as a Game developer and Game designer;

Technologies:

- C#;
- Back-end to Front-end relationship;
- Code organization;
- Work as a team in a short and stressful period;

Feb 2019 – Aug 2020

Undergraduate research: FUNCAMP (Unicamp Development Foundation)

Department of Systems and Energy DSE of the Faculty of Electrical and Computer Engineering (FEEC / UNICAMP).

Research title: Methodology for analyzing the presence of devices with Wi-Fi to estimate the passenger demand for Unicamp's internal circular buses.

Advised by Dr. Madson Cortes de Almeida.

Experiences:

- Developed a system using Wi-Fi signals to estimate passenger boarding, travel, and alighting, creating an origin-destination matrix for internal bus routes.
- Implemented software on ESP32 microcontrollers to detect smartphones, providing real-time passenger counts without compromising privacy.
- Enhanced data collection for transport planning, helping optimize routes and improve resource allocation on campus.

Technologies:

- Arduino;

- Embedded systems;
- Field research;
- Communication Wi-Fi;

Jul 2018 – Mar 2019

Technical capacitation – TT1: FAPESP (The São Paulo Research Foundation)

Company Mercúrio Automation Ltda. - ME/MERCURIO

Research Title: Feasibility analysis for an offline mobile payment system for vending machines.

Advised by Marcello Rodrigues de Oliveira.

Experiences:

- Developed a vending machine simulator in C/C++ for automated security testing.
- Created an Android application to evaluate the feasibility of transmitting encrypted data via Bluetooth.
- Tested Bluetooth signal strength and response time between nearby and distant devices.
- Concluded that Bluetooth transmission was inefficient for large data volumes, impacting the solution's feasibility.

Technologies:

- Git;
- Android Studio;
- C;
- App Android;

Languages

Brazilian Portuguese

Fluent (Native speaker)

English

Fluent

Personal projects

I've developed a simple website from scratch that summarizes my projects, providing easy access to my professional and academic CV, LinkedIn profile, and GitHub profile. For more information: viniciuscivali.com.