

# CARLOS YALTA VARGAS

Data Engineer & Applied AI | Data & Master Data Management | IT Project Management

carlos@cyv.me | linkedin.com/in/cyaltavargas | cyv.me | github.com/carlosyv

---

## PROFESSIONAL SUMMARY

---

Data engineer and analyst with over five years of experience across data management, ETL development, and system administration on Google Cloud, AWS, and Azure/Databricks. Background in economics and engineering, with hands-on exposure to applied AI, from supporting Databricks AI model Serving Endpoints in production to independent research on AI adoption and economic productivity. Moreover, a master's degree in entrepreneurship and executive education in leadership and innovation from MIT. Experienced leading agile teams and IT projects across international, cross-border environments (Peru, Sweden, and remote teams in the Americas, India and Philippines). Passionate about using data to drive innovation and measurable business value.

## SKILLS

---

**Data & Cloud:** Databricks (Certified Data Engineer Associate), Google Cloud Platform (GCS, gsutil), AWS, Azure, Snowflake, Informatica Cloud, Stonebranch, ORSOFT, Microsoft SQL, MySQL, PostgreSQL, Grafana, Redis Cache (NoSQL), System Administration (Linux/Ubuntu).

**Programming & Analysis:** Python (pandas, numpy, scipy, Dask, psycpg2, Django, BeautifulSoup, requests), SQL, R, Sphinx

**AI & Applied AI:** Databricks Model Serving Endpoints, applied AI-economics research (patents/TFP), exploring LLM-assisted development workflows and Front-end vibe coding using closed and open-sourced AI models.

**Analytics Tools:** Excel (Expert), PowerPoint, Access, SPSS, Stata, MATLAB, Jupyter Notebook, R Markdown

**Professional:** Agile Project Management, Team Leadership, Business & Financial Analysis, Consulting & Advisory, Stakeholder Presentation.

## PROFESSIONAL EXPERIENCE

---

### Data Specialist Data Engineer — Baufest

Apr 2023 – Present

Remote — clients in USA, Canada, Mexico, Brazil, and Argentina

- Perform Master Data Management tasks and build workflows using ORSOFT and Python for clients across the USA, Canada, Mexico, Brazil, and Argentina.
- Provide maintenance and incident support for master data tables on MDS/MySQL servers.
- Support Databricks-hosted AI model Serving Endpoints, including scheduled Jobs and task orchestration.

### Vinter (Kaiko) —

May 2021 – Apr 2023

Remote and On-site (Stockholm Metropolitan Area, Sweden)

#### Senior Data Manager

Sep 2021 – Apr 2023

- Developed and implemented procedures and rules for effective data management and data sharing across teams.
- Oversaw staff use of data systems, ensured compliance with legal and company regulations, and monitored systems affecting analytical results.
- Assessed system performance and recommended software, hardware, and storage improvements; managed all incoming data files.
- Assisted in upholding digital security systems to protect sensitive information.

#### Data Manager

May 2021 – Sep 2021

- Validated time-series indexes and cross-checked calculations against external sources; cleaned data and corrected database inconsistencies using Python (pandas, numpy, psycpg2).
- Built ETL processes with Python (requests, Django) to extract data from public APIs; managed files via Google Cloud Storage and gsutil.
- Administered five Ubuntu production servers; built Grafana dashboards to track CPU and database performance.
- Led a small team and handled technical recruiting.

#### Back End Developer (concurrent)

May 2021 – Sep 2021

- Developed API and web application features; wrote unit tests for Django models and views.
- Authored API documentation using Sphinx.

#### **Data Analyst**

*Jan 2021 – Apr 2021*

- Performed time-series analysis in Python (pandas, scipy) for financial decision-making, reporting directly to the CEO and CTO; presented results via Jupyter Notebook and R Markdown.
- Built web scrapers for commodities and bitcoin data using BeautifulSoup, requests, and urllib.
- Calculated price indices for Gold ETFs using Python (pandas, numpy).

#### **Assistant Professor — Universidad Nacional de Ingeniería**

*Jun 2020 – Nov 2020*

*Lima, Peru*

- Taught an elective course on business plan development, covering business model design, market and industry analysis, marketing plan, financial analysis, and operations planning.
- Led workshops on soft skills, leadership, entrepreneurship, innovation, and sustainable development.

#### **IT Project Manager — Telefónica del Perú**

*Nov 2017 – Jul 2018*

*Lima, Peru*

- Led software development projects and new business initiatives to increase company revenue, including the "Family Plan" telecom initiative reducing per-member billing.
- Managed 24/7 IT operations supporting more than 20 million customers across calls, data, and SMS services.
- Coordinated planning across multiple suppliers and delivered daily updates to the Chief Data Officer on the management of 5 million inactive customer accounts, including contingency planning.
- Bridged the Peru-based team with an offshore team in India to resolve issues quickly in a fast-paced, high-pressure, international environment.
- Led a team of five professionals managing incidents during active project development.

#### **KPI Analyst and Business Controller — Indra**

*Sep 2016 – Oct 2017*

*Lima, Peru*

- Administered contracts valued at USD 7 million annually.
- Designed and implemented new KPIs for supplier performance, increasing profits by 50% by accurately reflecting real supplier performance on critical functions in under 3 hours.
- Conducted financial and service audits and delivered monthly KPI presentations to management using PowerPoint.

#### **Project Assistant — SVC Servicios Integrales**

*Jan 2014 – Mar 2015*

*Lima, Peru*

- Coordinated with suppliers and assisted the logistics department during project execution, including facilities coordination for projects outside the capital.
- Performed cost evaluation and analysis for each project; prepared Excel reports for management.

### **SELECTED PROJECTS**

---

#### **World Bank Data Browser — [github.com/carlosyv/wb\\_data\\_app](https://github.com/carlosyv/wb_data_app)**

- Built and deployed an ETL pipeline and dashboard for the World Bank Indicators API, syncing the full catalogue into PostgreSQL and serving a web UI for browsing, downloading, and querying time-series data.
- **Stack:** Python 3.12, FastAPI, SQLAlchemy (async), Alembic, PostgreSQL 16, Docker Compose, GitHub Actions CI.

#### **AI Productivity Puzzle in Latin America — [github.com/carlosyv/tfp-ai-patent-latam](https://github.com/carlosyv/tfp-ai-patent-latam)**

- Independent research quantifying the relationship between AI patent activity and Total Factor Productivity growth across 9–17 Latin American countries (2000–2024), using Solow residual and DEA-Malmquist TFP measures.
- Implemented all panel-data econometric estimators (Pooled OLS, Two-way Fixed Effects, Random Effects, Common Correlated Effects) from scratch in NumPy/SciPy rather than off-the-shelf packages.
- **Stack:** Python, NumPy, SciPy, pandas.

## EDUCATION

---

### Online Program, Leadership and Innovation — MIT Professional Education

2020

- Santander Scholarship holder — 5-week program on AI, Blockchain, Cloud Computing, Cybersecurity, IoT, and Leadership.

### Master of Social Science, Entrepreneurship/Entrepreneurial Studies — Uppsala University

2018 – 2019

- Specialization in corporate entrepreneurship; key courses: opportunity recognition and development, business plan catalyst, managing the growing firm, strategic entrepreneurship, corporate entrepreneurship.

### 3S Shanghai Summer School — Shanghai University

2016

- Hanban Scholarship holder — 4-week program on Chinese language, culture, history, and economy.

### Engineer's Degree, Economics — Universidad Nacional de Ingeniería

2011 – 2015

## CERTIFICATIONS

---

- Databricks Certified Data Engineer Associate
- Databricks Academy Accreditation — Fundamentals
- Databricks Academy Accreditation — Generative AI Fundamentals
- Udemy certificates:
  - Relational Database Design
  - EssentialSQL: Data Modeling & Relational Data Architecture
  - Intro to Data Engineering
  - Building a Modern Data Warehouse – Data Engineering Bootcamp
  - Introduction to the Mongo DB
- Agile Project Management: Foundations
- Elements of AI

## LANGUAGES

---

Spanish (Native) · English (Full Professional) · French (Professional Working) · Swedish (Elementary) · Chinese (Elementary)

## PUBLICATIONS

---

- "Innovation Gene: The Relationship between Financial Performance and Communicating Innovation in Mission Statements."
- "Análisis de los determinantes de tiempo de operación de las MYPES" (2013).