

# Sebastian Sepulveda

[sbsepul.github.io](https://sbsepul.github.io) | [sbsepul@gmail.com](mailto:sbsepul@gmail.com) | [linkedin.com/in/sebasepulveda](https://linkedin.com/in/sebasepulveda) | [github.com/sbsepul](https://github.com/sbsepul)

## EDUCATION

---

### University of Chile

*Engineering in Computer Science*

Santiago

2017 – 2023

### University of Chile

*M.Sc. Data Science*

Santiago

2023 – 2024

Teaching aide in *Information Visualization (Autumn 2022)* *Design and Programming Methodologies (Fall 2020)*,  
*Algorithms and Data Structures (Spring 2019)*

## WORK EXPERIENCE

---

### Data Engineer

*Falabella Retail S.A*

Aug. 2022 – Present

*Santiago, Chile*

- Led the data migration for corporate planning processes across Chile, Colombia, and Peru, leveraging BigQuery for queries, Cloud Composer and Airflow for orchestration, Cloud Functions for data transformation, and other Google Cloud Platform tools.
- Automated and optimized data ingestion workflows using Python, enhancing efficiency across company.
- Developed business intelligence visualizations with Data Studio and Looker, enabling data-driven decision-making.
- Designed and implemented a new data transformation pipeline using DBT, replacing an Airflow-only workflow for improved maintainability.
- Established a scalable data infrastructure with separate development and production environments, using Terraform for infrastructure as code.

### Junior Backend Developer

*Business News Americas*

Jan. 2022 – Feb. 2022

*Santiago, Chile*

- Contributed to the development of an advanced search module for filtering news by text, date, and other attributes, improving client access to relevant content.
- Implemented a high-performance search system using Elasticsearch and Java, optimizing queries for large-scale document retrieval and ensuring scalability.

### Data Engineer Intern

*Seguros Falabella*

Dec. 2020 – Mar. 2021

*Santiago, Chile*

- Supported the migration of on-premise Oracle data to Google Cloud Platform (GCP), improving analyst accessibility.
- Refactored automated ETL processes from Pentaho to SQL-based pipelines in BigQuery, increasing data processing efficiency.
- Developed a standardized methodology for SQL-based ETL workflows in BigQuery, enabling internal teams to create and maintain processes autonomously.

### Web Scraping Developer

*Simple S.A*

Aug. 2020 – Oct. 2020

*Santiago, Chile*

- Developed a Python-based web scraping tool (Scrapy) to extract historical news data (2010–present) from Chilean media sources for a Spanish NLP study.
- Automated data extraction and storage using AWS EC2, MySQL and Python streamlining data collection for research purposes.

### Software Developer Intern

*Finantech*

Jan. 2020 – Feb. 2020

*Santiago, Chile*

- Researched web scraping methodologies for extracting financial data, including tables, text, and links from various sources.
- Developed a Ruby-based web scraping tool to extract and update financial company information, filling missing or outdated records.

## PROJECTS

---

**Tukuu** | *Node.js, Koa, AWS*

Aug. 2021 – Dic. 2021

- Collaborated in the develop of a backend web application using AWS (Lambda, ApiGateway, DynamoDB), Node JS and Koa framework serving a REST API with React as the frontend

**Functional Diversity Support Unit, University of Chile** | *Django REST, Docker, Git* Mar 2021 – Jun. 2021

- Developed a backend web application using Django REST serving a API with React as the frontend.
- Collaborated in the design of the data model using django to create and modify tables.

## SKILLS

---

### Technical

**Languages:** Python, Java, C/C++, SQL, Ruby, (Postgres, BigQuery, MySQL), JavaScript, HTML/CSS, PHP, R

**Frameworks:** Node.js, Django Rest

**Clouds:** AWS, GCP

**Developer Tools:** Git, Docker, Google Cloud Platform, Amazon Web Services, Tableau

### Machine learning and Artificial Intelligence

**Computer Vision:** OpenCV, Detectron2, MMDetection, YOLOv8-v5 by Ultralytics, albumentation, Pytorch, Keras (Vision models), Hugging-face (transformers)

**Machine Learning:** Sklearn, Seaborn, matplotlib, pandas, numpy.

**NLP:** torchtext, gensim.

### Language

**English:** TOEFL ITP B1

**French:** DELF B1