Sandro Loch

es.loch@gmail.com | linkedin.com/in/esloch | github.com/esloch

55-47-98455-6932

Santa Catarina - Brazil

Summary

I am a Full-Stack DevOps Developer with over five years of experience designing and implementing scalable backend systems, automating infrastructure, and contributing to impactful open-source projects. My work focuses on developing robust solutions for data-intensive applications, ensuring high availability, performance, and automation throughout the development lifecycle.

Experience (last 5 years)

Software Engineer, OSL - Open Science Labs / LiteRev - Université de Genève 12/2023 – present

- Developed and optimized backend architecture using Docker, NGINX, Django, Celery, Redis, and Elasticsearch for automating scientific data collection and processing.
- Developed and integrated Retrieval-Augmented Generation (RAG) workflows to improve AI-driven document retrieval.
- Implemented CI/CD pipelines with GitHub Actions for seamless deployment and automated testing.
- Managed SSL certificates and configured NGINX servers for secure and efficient access.
- Developed automated pipelines for collecting and processing scientific articles from sources like arXiv, MedRxiv, and PubMed Central using APIs and Django.
- Integrated Elasticsearch for efficient data indexing and retrieval, enhancing search performance.
- Led migration and configuration of virtual servers, ensuring optimal PostgreSQL deployment.

Full-Stack Backend Developer, Fiotec - Scientific and Technological Development in Health Foundation 03/2019 – 12/2024

- Developed and maintained Django applications for reporting dengue and other arboviruses cases.
- Created and maintained PostgreSQL databases, including modeling, massive data loading, and query optimization.
- Developed APIs with Django Rest Framework for querying epidemiological data.
- Optimized the development pipeline using Ansible, Docker, Celery, PostgreSQL, and CI/CD (GitHub Actions).
- Managed large volumes of data with Pandas and Dask, producing interactive reports and analyses. Automated climate data ingestion by integrating APIs and processing data using Python, Pandas, and SQLAlchemy.
- Implemented PostgreSQL replication using Bucardo, ensuring high availability of epidemiological data.
- Created scripts and playbooks to synchronize and validate historical alerts in PostgreSQL with Ansible automation.
- Deployed scalable object storage solutions with MinIO and Docker for epidemiological data repositories.

Technologies

- **Programming and Scripting:** Python, SQL, Bash
- **Web Development and APIs:** Django, Django Rest Framework, Django ORM, Bootstrap
- **Data Management and Analysis:** PostgreSQL, SQLAlchemy, Bucardo, MinIO, Pandas, Dask, Ibis
- **DevOps and Automation:** Docker, Ansible, GitHub Actions, Celery, APScheduler
- **Networking and Infrastructure:** Linux Servers, DNS, DHCP, Group Policy, Hybrid PABX Systems
- **Visualization and Reporting:** Plotly, Matplotlib, Highcharts
- **Package Management:** Conda, Poetry, Pip
- **CI/CD and Testing:** GitHub Actions, deployment pipelines, unit testing
- **Tools and Technologies:** Elasticsearch, Redis, NGINX, Let's Encrypt
- **Security Compliance:** ITIL best practices, data integrity, and process automation

Courses Certifications

- **MySQL Explorer;** Oracle University (Sep 19, 2024)
- **Scrum Foundation Professional Certification - SFPC™;** CertiProf® Professional Knowledge (July 02, 2023)
- **Data Analysis for Health Surveillance;** UFSC: Federal University of Santa Catarina (May 15, 2023)
- **SGDB PostgreSQL;** IFRS: Federal Institute of Rio Grande do Sul (Apr 02, 2022)
- **Leadership Coaching;** Leading for High-Performance (SEBRAE) (Nov 12, 2021)
- **Statistical Foundations;** LabTime (CAPES) (Nov 11, 2020)
- **Introduction to Data Visualization with Matplotlib;** DataCamp (Sep 16, 2020)
- **Visualizing Geospatial Data in Python;** DataCamp (Sep 08, 2020)
- **Unit Testing for Data Science in Python;** DataCamp (Jun 10,2020)
- **Intermediate Importing Data in Python;** DataCamp (May 28, 2020)
- **Introduction to SQL;** DataCamp (May 22, 2020)
- **Introduction to Importing Data in Python;** DataCamp (May 13, 2020)
- **Bash Scripting;** DataCamp (May 20, 2020)
- **Conda Essentials;** DataCamp (May 04, 2020)
- **Git;** DataCamp (May 01, 2020)
- **Manipulating DataFrames with Pandas;** DataCamp (Jan 14, 2020)
- **CCNA; Network fundamentals, Network access, IP connectivity, IP services, Security fundamentals, Network Protocols, Automation and programmability;** SENAC/CISCO (2006)
- **Windows 20003Server / Concepts and Infrastructure Network / Infraestruture and Active Directory;** TREITEC (2004)
- **Delphi 6 Basic;** ProWay Informática (Mar 08, 2003)

Projects and Open Source Contributions

- **SecondDx:** Contributed to the development of a platform that enhances diagnostic accuracy through AI-powered second opinions, focusing on patient anamnesis, medical imaging, and health prognostics.
- **MyMHAI - Mental Health Care:** Contributed to the development of a mental health platform focused on leveraging AI technologies to support mental health assessment and interventions. Developed data-driven tools for analyzing mental health patterns, integrated AI models for predictive analytics, and enhanced the accessibility of mental health resources through automated systems.
- **Literev (TheGraphNetwork):** Developed tools for scientific literature review, focusing on Retrieval-Augmented Generation (RAG) workflows for automated document retrieval and knowledge extraction. Contributed to the RAGO library, optimizing document retrieval and AI-assisted knowledge extraction with OpenAI models. Additionally, supported other data-driven projects within TheGraphNetwork, enhancing public health and scientific research applications.
- **InfoDengue:** Contributed to the development of data-driven tools for monitoring arboviruses in Brazil, focusing on epidemiological data analysis and predictive modeling. Supported the AlertaDengue web portal for real-time monitoring of arboviruses by integrating climate, health, and social data. Developed PySUS, a Python library for accessing and analyzing datasets from Brazil's Unified Health System (SUS), and automated climate data collection through the CrawlClima tool to support public health models. Implemented scalable storage solutions with MinIO (StorageBox-MinIO) for managing epidemiological data and designed PostgreSQL replication solutions using Bucardo (infodengue-bucardo) to ensure data consistency and high availability.
- **EpiGraphHub (TheGraphNetwork):** Contributed to epidemiological data platforms for public health analysis, focusing on data integration, visualization, and real-time analytics.
- **OpenScienceLabs:** Active contributor to open-source projects promoting data accessibility, reproducibility, and scientific research tools, supporting initiatives like Literev and other innovative data-driven applications.

Education

University Leonardo Da Vinci, Santa Catarina, BR, Big Data and Analytical Intelligence — Sept 2021 – Dec 2024

World Quant University (WQU), New Orleans, LA, Data Science Lab — Dec 2024 - In Progress