João Correia

jpcorreia99@gmail.com | linkedin.com/in/joao-cor | github.com/jpcorreia99

Experience

Palantir Technologies March 2023 – Present
Developer - Backend | Java, Python, Rust *London, UK*

- Reduced by ∼**8x** python environment build times for python pipelines. (Announcement.)
- Reduced by ∼**9x** startup performance of two crucial services. (Announcement.)
- Owned Data Expectations, a python library for pipeline reliability checks.
- Developed Hawk, a python package manager **20%** faster than SOA alternatives.
- Mentored Intern and New Grad engineers, leading to successful integration and conversions.
- Maintained critical data integration systems, improving telemetry via real time automated monitoring.

Amazon Web Services (AWS) Aug. 2022 – Jan. 2023
Software Development Engineer Intern | Java, Cloud *Dublin, Ireland*

- Engineered migration of the internal monitoring system to native AWS Cloudwatch.
- Wrote design documents for features on the migration critical path.
- Reduced database occupancy by 30% by identifying stale and deprecated entries.

Calouste Gulbenkian Foundation Otc. 2020 – Sept. 2021
Artificial Intelligence Researcher | Python, Tensorflow *Remote, Portugal*

- Developed new text *LSTM* based classification algorithms for *European Portuguese*.
- Topped existing systems accuracy by **over 20%**.
- Publication accepted in IDEAL'21.

Subvisual July. 2020 – Oct. 2020
Software Engineer Intern | Ruby on Rails, Web dev *Braga, Portugal*

- Developed web app for managing the monthly gatherings of the book club at the company.
- Open source project written in Ruby on Rails, integration with Google Books' API.

Education

École polytechnique fédérale de Lausanne (EPFL) Lausanne, Switzerland
Msc. in Computer Science *Sept. 2021 – Feb 2023*

- Grade: **5.3/6**.
- Specialization in distributed systems.

University of Minho Braga, Portugal
Bsc. in Software engineering *Sept. 2018 – Aug 2021*

- Final grade: **17.3/20**. Top Student: **Ranked 1/200**.

Projects and Significant Open Source contributions

Apache Spark (Contribution) | *Scala*

- Allowing users to control Kubernetes memory allocation for Spark nodes.
- Announced as new feature in Spark 4.0.

Rattler (Contribution) | *Rust, Conda*

- Solved critical bug for streaming zip files compressed via data descriptors.

Safira (Project) | *Elixir, Pheonix*

- Backend for event management built with Elixir for high scalability.
- Leverages Phoenix's actor model to handle thousands of concurrent queries.

Peerster (Project) | *Go, Blockchain*

- Decentralized Gossip-based P2P system.
- Implements multiple classical algorithms such as BitTorrent, Paxos and a blockchain ledger.