Pablo Panero

Software Engineer

(+34) 622 240 450
/in/ppanero
ppanero
ppanero.github.io
pablopanerovz@gmail.com

Technical Skills

Programming: Python, JavaScript, Go

(Web) Technologies: Flask, Celery, SQLAlchemy, Pytest, React, Redux, Webpack, Semantic-UI, Jinja

Databases: Elasticsearch, PostgreSQL, Redis

(Big) Data Science: Spark, Kafka, Flume, YARN, Tensorflow, Keras

Other: Docker, OpenShift, Helm, OpenStack, Puppet, Fabric, GitHub actions, Kibana, Grafana

Personal skills: Team player, Problem solving, Initiative, Organization, Curiosity driven

Languages

Spanish: Native language

English: Full working proficiency

French: Full working proficiency

Education

MSc. Data Science
Universidad Oberta de Catalunya
2018 - 2021
Catalonia, Spain

Erasmus+ Exchange
Oulun Yliopisto
2014 - 2015
Oulu, Finland

BSc. Computer Science Engineering
Universidad Complutense de Madrid
2011 - 2016
Madrid, Spain
Graduated with extra credits

Experience

Software Engineer | CERN - Zenodo, InvenioRDM — Jun 2019 - Present

Python, Celery, Elasticsearch, Docker, OpenShift, Helm, RabbitMQ

Core developer of Zenodo and InvenioRDM:

- **Maintain** the Invenio Framework (GitHub), with **more than 130 modules**.
- **Design a migration strategy for over 3M documents and 10M files (1.7PB).**
- Redesign and implement core packages to **build RESTful APIs, data models**, manage persistent identifiers and support event handling.
- Develop Helm charts to automate the deployment on OpenShift.
- **Operations and incident response on Zenodo, a service with 25M visits/year and 350k active users**.
- Research the usage of neural networks on SPAM detection.
- **Scrum master** of sprints of 6 software engineers.
- Gather requirement from stakeholders, as well as user feedback sessions.
- Assist in the organization of workshops of more than 35 people (21 partner institutions).
- Participate on outreach events and present webinars.
- Lead the *buddy system*, for newcomers onboarding.
- **Lead a collaboration working group** to identify and tackle problems regarding remote working practices. Define guidelines and success metrics. The results had an impact on more than 15 people.

Software Engineer | CERN - CERN Search — Feb 2018 - Jun 2019

Python, React, Elasticsearch, Docker, OpenShift, RabbitMQ

Lead developer of the new generation CERN Search as a Service:

- From design to **production in less two years time** (*one man project*).
- Maintain a production instance that indexes and searches **over 24M documents**. Including the optimization of Elasticsearch mappings and search scoring functions.
- Provide support to clients (e.g. Indico) on their integration with the service.
- Develop and maintain the crawler used to harvest the whole CERN websphere.

Big Data Engineer | CERN - Computer Security — Feb 2017 - Feb 2018

Python, Scala, PHP, Spark, Kafka, Flume, YARN, Mesos, Puppet, Docker

Core developer of the Compute Security Team's infrastructure:

- Develop a **large scale intrusion detection system** capable of processing 500GB of data per day (50k events/s).
- Contribute new features to the Malware Information Sharing Platform (**MISP**). **Deploy and maintain two production instances** (CERN and WLCG).
- Develop and maintain the MISP puppet module on **Vox Pupuli**.

Big Data Instructor | Culture Lab and NewProLab — Nov 2016 - Nov 2017

Hadoop, Spark, Kafka, Flume, Hive, Pig, Impala, Sqoop, Cloudera, Hortonworks, YARN

- Design and teach both in-person and online Big Data related courses.