Antonio Ercole De Luca

GITHUB https://github.com/eracle
SKYPE ercole.deluca •
E-MAIL eracle@posteo.eu

Data Engineer and Software Developer with experience on Python and Java from 2013.

Mainly focused on all aspects of data management such as Retrieval, Storage, Scaling, Analysis and Mining. Usual data flow is Web Scraping, Relational Database, Cloud infrastructure, REST interface or Jupiter notebook.

Always fashinated by the potential given intersecting data coming from the Web and Data Mining techniques such as Data Clustering, Graph Analysis, Information Retrieval and Machine Learning.

Github Projects:

linkedin: Scraping software aimed to collect info from linkedin users. It uses scrapy, docker, Selenium Webdriver and Python3.

Some Actual Skills:

Python3, Java, Groovy, C, Bash;
Git, PyCharm, Gradle;
docker, docker-compose, Amazon AWS, heroku, Ubuntu Server;
Django, DRF, Scrapy, Selenium Webdriver;
postgreSQL, NEO4J, elasticsearch, celery, redis;
CrowdFlower; BPMN, Petri Nets; TDD;

Working Experience:

2019   **ErnieApp**
Dublin, Ireland
Remote
Backend Test Developer
python3, docker, django, celery, postgres, on Heroku.

2018	**XPLOD Social** **Los Angeles, California** **Remote** Software Architect

Design of an Analytics Platform to show insights on data
gathered from Instagram;
python3, docker, django, celery, redis,
Selenium, Instapy, Laravel, Mysql, on DigitalOcean.

2018	**Trademark Pro** **Canada** **Remote** Data Scientist

Image similarity search engine (MVP);
python3, docker, flask, keras, tensorflow,
scipy, numpy, annoy, postgres, on EC2 and S3.

2018	**Kiwicampus.com** **Berkeley, California** **On site** Backend Software Developer

In charge to build the Integration platform with Doordash Inc.
django, postgres, celery, redis, postgres, on Heroku.

2018	**Decyde.io** **San Francisco, California** **Remote** Python Software Developer

python, scrapy, scrapinghub.com, google-custom-search,
selenium**.**

2017	**Einhorn Manufaktur** **Berlin, Germany** **Remote** Backend Software Developer

Backend REST APIs Developer.
Python3, Django, Postgres, TDD.

2016   **Legalkite Ltd**
             Zurich, Switzerland
             Remote
             Backend Software Developer

             REST APIs, Data storage and Cloud infrastructure.

             Python3, Django, Amazon AWS, ElasticSearch, Agile, TDD, Jira.

2015   **SpazioDati s.r.l.**
             Pisa, Italy
             On site
             Internship/Thesis

             Supervisors: Prof. Paolo Ferragina and Ugo Scaiella

             Development of an ETL workflow which involved CrowdFlower
             (now Figure Eight). Particular focus was given to cost efficiency
             by leveraging an Active Learning technique.

             The output dataset was used to train a logistic regression
             model that replaced one of the steps of SpazioDati's proprietary
             Entity Linking system (SaaS): namely Dandelion

             Tools used: Groovy, Intellij IDEA, WEKA, Gradle, Git and Ubuntu
             server.

2013   **Synthema S.R.L.**
             Pisa, Italy
             On site
             Software Engineer

             Design and implementation of Java modules for Crawling and
             Linguistic Analysis (Romanian, Japanese, Chinese and German),
             as part of the CAPER european project.

Education:

2006   **Bachelor Degree**
             Laurea triennale
             Computer Science
             Informatica
             University of Pisa, Italy
             Outstanding final marks 110/110

Thesis: Document classification using lexical and semantic resources.
(Catalogazione di libri tramite risorse di semantica lessicale)

2009   **Erasmus LLP Program**
Facultad de Informática
Universidad Politécnica de Valencia, Spain

2011   **Postgraduate Master's Degree**
Laurea Magistrale
Business Informatics
Informatica per l'Economia e l'Azienda
University of Pisa, Italy
Outstanding final marks 110/110
Thesis: Applying Active Learning on Entity Linking.
(Algoritmi di Active Learning applicati al problema dell'Entity Linking Pruning).

2019   **Bachelor's Degree**
Laurea triennale
Statistics for Data Analysis
Statistica per l'analisi dei dati
Università di Palermo, Italy

Publications:

Semantic Crawling: an Approach based on Named Entity Recognition
Proceedings of FOSINT-SI 2014
Antonio Ercole De Luca, Carlo Aliprandi, Giulia Di Pietro, Matteo Raffaelli, Tiziana Soru.

CAPER: Crawling and Analysing Facebook for Intelligence Purposes
Proceedings of ASONAM2014
Carlo Aliprandi, Antonio E. De Luca, Davide Gazzè, Mariantonietta N. La Polla, Giulia Di Pietro, Matteo Raffaelli, Andrea Marchetti, Maurizio Tesconi.

Languages:

English: Fluent
Spanish: Fluent
Italian: Native

Antonio Ercole De Luca
GITHUB https://github.com/eracle **SKYPE** eracole.deluca **E-MAIL** eracle@posteo.eu