

Antonio Ercole De Luca

Data Architect

Jr. Quantitative Developer

PROFILE

I am a Data Architect, working as a consultant since 2016, having shifted over different roles such as Software Developer, Data Scientist, ML Engineer, DevOps and Scrum Master.

I have worked on all aspects of data management such as Retrieval, Storage, Cloud Scaling, Analysis and Mining; encompassing Web Scraping, Design of Relational Databases, Deployment of Cloud Infrastructure, RESTful interfaces and Jupiter Notebooks.

In 2019 I left consultancy to become a quantitative developer.

During my career, I also contributed on two scientific papers on the field of Web Scraping and Data Visualization.

SKILLS

+ Python	<div><div></div></div>
+ SQL	<div><div></div></div>
+ Pandas	<div><div></div></div>
+ Django	<div><div></div></div>
+ Docker	<div><div></div></div>
+ Kubernetes	<div><div></div></div>
+ Regression	<div><div></div></div>

LANGUAGES

+ Italian	<div><div></div></div>
+ English	<div><div></div></div>
+ Spanish	<div><div></div></div>

WORK EXPERIENCE

Consultant, Remote
(2017 - 2019)

Data Architect

Worked with more than 10 international clients on Upwork.com receiving 5/5 stars as median feedback.

Some stories:

- Designed an image similarity search engine with near zero response time, indexing more than 40 000 imgs by using Approximate Nearest Neighbors data structure;
- Reduced client's cloud infrastructure costs by dockerizing a django app and migrating from Digital Ocean to Google Cloud;
- Built a django app with 83% testing coverage by using Test Driven Development.

Legalkite Ltd, Zurich
(2016 - 2017)

Back-end Software Engineer

Design of several RESTful endpoints;
Improved 10x REST API response time by translating django ORM queries into raw SQL queries;

Solved data inconsistency bugs by:

- Refactoring postgres schema;
- Making Elasticsearch a read-replica;

Scaled 100x scraping module by porting it to scrapy and scrapinghub.

SpazioDati, Pisa, Italy
(2015-2016)

Machine Learning Engineer

Design of a Human-in-the-loop Active Learning workflow using CrowdFlower API (now Appen).
Increased 2x F1-score in Logistic Regression model used for Named Entity Linking by using Uncertainty Sampling.

Synthema, Pisa, Italy
(2014-2015)

Software Developer

Productionize Natural Language Processing steps: Segmentation, Tokenization, POS, Stop Words, Dep.Parsing, NER; (Romanian, Japanese, Chinese and German)
EU funded (5.6 Mln €) project: CAPER.

EDUCATION

Bachelors Degree
Statistics & Data Science
Università di Palermo, Italy
Financial Mathematics, Finance Theory, Multivariate Analysis, Inference, R Language, Probability. (2019 - 2020)

Masters Degree
Data Science and Business Informatics
University of Pisa, Italy
Thesis on Natural Language Processing:
Applying Active Learning on Entity Linking. (2011 - 2013)

Bachelors Degree
Computer Science
University of Pisa, Italy
Thesis on Natural Language Processing:
Document classification using lexical and semantic resources. (2007 - 2010)

REFERENCES

Massoud Zekeria
Founder at Legalkite Ltd
massoud.zekeria@gmail.com

Sasha Iatsena
Product Manager at Google
sasha@iatsenia.xyz

Ugo Scaiella
Software Development Manager
at SpazioDati
P. +39 050 754243
scaiella@spaziodati.eu



e-mail:
eracle@posteo.eu

Personal web site:
<https://eracle.me>