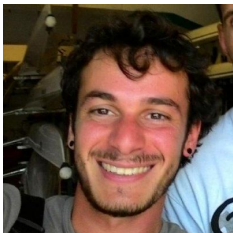


PERSONAL INFORMATION

Antonio Ercole De Luca

 Pisa, Italy eracle@posteo.eu Skype: ercole.deluca

Sex Male | Date of birth 04/11/1987 | Nationality Italian

JOB APPLIED FOR
POSITION
PREFERRED JOB
STUDIES APPLIED FOR

Data Science - Software Development

WORK EXPERIENCE

01/11/2016 – 31/07/2017
9 Months

Senior Software Architect

Legalkite Ltd

Remote

<https://www.legalkite.ch/>**Roles:**

Responsible of the REST APIs, Data storages and Cloud infrastructure.
Python, Django, Amazon AWS, ElasticSearch, Agile, TDD, Jira.

About Legalkite:

Legalkite is a fully remote Swiss start-up which provides a cloud based software aimed to change the way laws and legal content are consumed and explored. The software contains a full-text search experience over the Swiss laws, and allows the users to store and share legal know-how.

Business or sector Software Development

15/03/2016 – Today

**Freelance
Software Developer - Data Scientist**

Upwork

Remote

<https://www.upwork.com/>**Role:**

Senior Software Developer with 4 years of experience on Java, Python and related technologies.

Mainly focused on Web Technologies and Data Mining, with special interest on concepts such as Data Clustering, Graph Analysis, Information Retrieval, Machine Learning and Web Scraping.

Open Sourced Projects:

[gap](#): Data Mining library with special focus on Clustering of non-vector instances.(Java)

[fbot](#): Facebook automation library implemented with Selenium Webdriver. It can be used also for crawling/scraping all the posts in a Facebook group. (Python3)

[linkedin](#): Scraping software aimed to visit as more pages of the linkedin users as possible for the purpose of gaining visibility. It uses Scrapy, Selenium Webdriver, Firefox 45. (Python3)

[rbroker](#): Simple REST APIs, with Django REST Framework, that synchronizes remote mobile devices executing Tasks. (Python3)

Business or sector Software Development

9 Month:
1/06/2015 – 1/03/2016

Master's degree Thesis/Internship

SpazioDati s.r.l.

Polo Tecnologico di Navacchio – Pisa

<http://www.spaziodati.eu/en/>

Supervisors:

[Prof. Paolo Ferragina](#)

Doc. Ugo Scaiella



The main objective of the thesis was the development of a methodology for the automatic adaptation of a generic Entity Linking system to new (and possibly unseen) linguistic contexts.

The implemented process uses a web based Crowd-sourcing platform, namely [CrowdFlower](#), which allows to submit simple questions (known as micro-tasks) to a pool of distributed knowledge workers through the use of a REST API.

The work done on the thesis was performed in a 9 month internship at [SpazioDati](#), a young start-up based in Pisa and Trento, which develops software systems for analysis of Big Data, they have their own Entity Linking system offered as a SaaS product: known as Dandelion.

The methodologies proposed on the thesis were used during the internship to train a logistic regression model used to perform the Pruning phase of the entity linking process of [Dandelion](#).

In order to choose which new data instances to label (to minimize the labelling costs) an Active Learning algorithm was used: namely Uncertainty Sampling.

The tools used were: the programming language Groovy, IntelliJ IDEA, the WEKA library and its Logistic class, the build system Gradle, Git and Ubuntu server.

Business or sector Machine Learning

9 Month:
1/10/2013 – 30/06/2014

Software Engineer

Synthema S.R.L.

Via Malasoma, 24 – Ospedaletto - 56121 Pisa (PI) - ITALIA

+39.050.9656.411 info@synthema.it



Member of Synthema Team for the **CAPER** project:

The goal of the European project was to create a common platform for the prevention of organised crime through sharing, exploitation and analysis of Open and private information sources.

The CAPER project was supported by the European Commission through the Seventh Framework Programme for Research and Technological Development with up to 5.6 Mio €, out of a total budget of 7.1 Mi. €. The 3-year project has been running since 1st July 2011.

Software Engineer and Developer:

Design and implementation of several Java and PHP software modules used for Crawling and for Linguistic Analysis purposes. The modules were: some SOAP based back-ends (Web Services) and some Wrappers for external open-source Entity Recognizers which covered different languages (Romanian, Japanese, Chinese and German).

Linux Server Administrator:

Responsible for Deployment, Configuration, Management and Release of the private Linux Virtual Machines needed for the CAPER project. The OS distributions were CentOS 5 and Ubuntu Server 12.04, running upon an enterprise-private data center with VMware ESXi 5.5.

Natural Language Processing Researcher:

We published two research papers on two International Conferences: FOSINT-SI 2014 and ASONAM 2014. We presented: a Web Crawler with the ability of focusing only on certain types of Named Entities and a Graph visualization tool usable with the data extracted from the discussions of Facebook, they were both used in the CAPER project.

Business or sector Natural Language Processing

6 Month:
1/02/2011 – 30/06/2011

Bachelor's degree internship

Department of Computer Science - University of Pisa

Largo B. Pontecorvo, 3 , 56127 Pisa

Telephone: +39 050 2212700



Under the supervision of prof. Vincenzo Gervasi I spent a five months research period at the Department of Computer Science of University of Pisa. I developed a software which performs Semantic Text Tagging for the Italian language through the use of Wikipedia Italia as a semantic resource.

Business or sector Natural Language Processing

EDUCATION AND TRAINING

2011 - 2016

Postgraduate Master's Degree (Laurea Magistrale), Business Informatics (Informatica per l'Economia e l'Azienda)

EQF level: 7

University of Pisa, Pisa , Italy

The two year graduate program in [Business Informatics](#) has been designed to meet the growing demand for professionals with an interdisciplinary skill both in informatics and in business to satisfy the increasing demand by companies to compete using analytic methods. The graduate program is focused on Business Intelligence techniques to support decision making.

Computer Science Courses: Information Retrieval, Programming Internet Applications, Web Mining and Social Network Analysis, Data Mining, Natural Language Processing, Introduction to Logistics, Business Performance Analysis, Decision Support Information Systems.

Economics Courses: Business Administration II, Computer Science Law, Planning And Management Control, Business Management.

Thesis: Active Learning algorithms applied to the Entity Linking Pruning problem.
(Algoritmi di Active Learning applicati al problema dell'Entity Linking Pruning).

Final Mark: 110/110

2006 - 2011

Bachelor's Degree (Laurea triennale), Computer Science

EQF level: 6

University of Pisa, Pisa, Italy

The three-year undergraduate programme in Computer Science aims to integrate the comprehension of information technology and its application in problem solving activities with a solid scientific and methodological preparation.

Mathematics Courses: Mathematical Methods, Algebra, Mathematical Analysis, Physics, Numerical Calculus, Calculation Of Probability And Statistics, Operations Research, Simulation.

Programming Courses: Principles of Computer Programming, Methodology of Computer Programming, Recursive Programming, Algorithms, Languages.

Programming Laboratories: Introduction to Programming Laboratory, Structured Data Programming Laboratory, Network Programming Laboratory, Lab-based Course on Competing Programming Techniques and Systems, Systems Language Laboratory.

Computer Science Courses: Databases, Software Engineering, Operating Systems, Computer Networks, Computer Architectures, Computer Science and Culture, Intelligent Systems I, Neural Networks I.

Languages Courses: English, Spanish.

Thesis: Document classification through the use of lexical and semantic resources.
(Catalogazione di libri tramite risorse di semantica lessicale)

Final Mark: 110/110

2008 - 2009

Erasmus LLP Program

Facultad de Informática - Universidad Politécnica de Valencia, Spain

Spanish language certificates: A2, B1.

PERSONAL SKILLS

Mother tongue(s) Italian

Other language(s)

	UNDERSTANDING		SPEAKING		WRITING
	Listening	Reading	Spoken interaction	Spoken production	
English	C1	C1	C1	C1	C1
Spanish	C1	C1	C1	C1	C1

Levels: A1/2: Basic user - B1/2: Independent user - C1/2 Proficient user
Common European Framework of Reference for Languages

Job-related skills

Specialties:

Programming Languages: Python3, Java, Groovy, C, Bash;

Tools: git, github, pip, Pycharm, Gradle, Maven, Git, Junit, Eclipse, IntelliJ IDEA;

Operating Systems: Ubuntu Server, CentOS, VMWare ESXi 5.5;

Frameworks (Software Dev.): Django, Scrapy, Selenium Webdriver, Tomcat, Axis2, CXF, Apache2, Jung 2.1;

Frameworks (Data Mining): Weka, Knime;

Technologies: REST and RESTful, JSON, XPATH, XML, WSDL, SOAP;

Databases: SQLite, MySQL, Postgres, ElasticSearch;

Data Enrichment Platforms: CrowdFlower;

Abstract Modelling: UML, BPMN, Petri Nets;

Methodologies: Scrum, Agile, TDD.

Other skills Drummer and Percussionist

Driving licence B

ADDITIONAL INFORMATION

Publications [Semantic Crawling: an Approach based on Named Entity Recognition](#)

Proceedings of FOSINT-SI 2014

Authors: Antonio Ercole De Luca, Carlo Aliprandi, Giulia Di Pietro, Matteo Raffaelli, Tiziana Soru.

CAPER: Crawling and Analysing Facebook for Intelligence Purposes

Proceedings of ASONAM2014

Authors: Carlo Aliprandi, Antonio E. De Luca, Davide Gazzè, Mariantonietta N. La Polla, Giulia Di Pietro, Matteo Raffaelli, Andrea Marchetti, Maurizio Tesconi.