
Backend Developer / Data Scientist

Functional Programing, Information Retrieval, Machine Learning, Semantic Web, Big Data.

Technical Skills

Languages: **Clojure, Python, Java**, Scala, Matlab, Haskell, Javascript.

Databases: **Elasticsearch, MongoDB, Kafka, Virtuoso, Druid**, Redis, MySQL, PostgreSQL.

Semantic Web: **RDF, SPARQL, RDFs, OWL**, Ontologies.

Machine Learning: **Classification, Regression, Social Network Analysis**, Anomaly Detection, Recommender Systems.

Frameworks: **Scikit-learn, Lucene, Gephi, Docker**, Mesos/Marathon.

Professional Experience

KCityLabs: Backend developer / Data scientist (2017 - now)

Developing an API that controls the whole life cycle of a geographical knowledge base, which links more than 20 data sources. Building a regression model to predict dwelling prices on french riviera. This model is trained on more than 100k real estate transactions, taking benefits not only of dwelling features but also of their geographical environment.

Technologies: Python, Clojure, MongoDB, MySql, Machine Learning.

Vigiglobe: R&D Engineer / Backend developer (2014 - 2017)

Developed a real time social media analytics platform which mines and aggregates thousands of messages per second. My contributions range from backend development (80%) to research on NLP and Social Media mining (20%). In particular, I have developed components for brand recognition, geolocation, search, and statistical trend detection.

Technologies: Clojure, Python, Semantic Web, Elasticsearch, MongoDB, Virtuoso, Druid, Redis, Kafka, Lucene.

Mnemotix: R&D Engineer / Semantic Web developer (2013 - 2014)

Developed a distributed medical knowledge base for e-ginseng ANR project enabling large-scale epidemiology analysis on distributed autonomous medical data sources.

Technologies: Java, Semantic Web, Gephi Toolkit, MongoDB.

Alten: Semantic Web Developer at Orange (2011 - 2013)

Developed a knowledge base, taking advantage of Semantic Web technologies, to supply business intelligence frameworks and information retrieval services for Orange search engine.

Technologies: Java, Python, Semantic Web, Virtuoso, Dbpedia.

Orange Labs / INRIA: Research Engineer (2007 - 2010)

Ph.D. thesis on semantic social network analysis.

<https://tel.archives-ouvertes.fr/tel-00586677/>

Technologies: Java, Semantic Web, Gephi.

Education

2011:

Ph.D. in Computer Science, Telecom ParisTech, France.

2007: Master degree in Computer Science, Polytech Nice Sophia, France.

Some Publications

2014: Global Initiative for Sentinel e-Health Network on Grid (GINSENG): Medical Data Integration and Semantic Developments for Epidemiology. S. Capiere, G. Éréteo et al, 14th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing

2011:

Semantic Social Network Analysis.

G. Ereteo, Ph.D. Thesis.

2009:

Analysis of a real online social network using Semantic Web frameworks.

G. Éréteo et al, 8th International Semantic Web Conference.

Certifications

2017: Machine Learning by Stanford University on Coursera.

2016: PSC1 (first aid level 1).

2007: TOIC score 825