

ANTONY BERNADOU

Big Data Scientist - Looking for a short term mission in Data Analytics

 +61 490 03 44 8

 antony.bernadou@gmail.com

 www.linkedin.com/in/antonybernadou

Personal Profile

I am a French engineer living in Sydney for both personal and professional development. I have full rights to work in Australia, up to 6 months for the same company.

Graduated in Mathematics and Computer Science, I am truly passionate about Analytics and Data Science. Since 2015, I have been working on various data-oriented projects, mainly on top of Hadoop ecosystems. As a consultant, I have worked on different environments, which have enabled me to understand how to succeed in implementing Big Data projects from conception right through to delivery and acceptance.

Being proactive and curious helps me to keep updated of recent Frameworks and tools used by Data Scientists or Data Engineers. I have been in charge of a Data Lab for my last company, Sogeti, it was an occasion to share knowledge and skills on Big Data, IA, Cognitive Science and other Analytic subjects.

Career Accomplishment

Speaker at a Big Data event in Bordeaux (France), providing possible answers to following questions:

- How to easily deliver Machine Learning model designed by Data Scientists?
 - How to give the best chance to your Data Scientists by creating a Data Lab?
 - What do Agile methodologies and tools offer to Big Data applications?
-

Professional Experience

Big Data Engineer Ext. Consultant, Air France KLM — 2016-2017

As a Big Data consultant for Air France KLM in Customer Data Management service, I have been working on the following projects:

- Booking Process Data Unification:

Responsible for the architecture and the design of a Big Data application. Scrum master role: help the development team to adopt priorities according to business needs evolution, negotiation with Product Owner and technical support. Handle more than 5 Tb of Data,

Technical environment: HDP 2.6, Spark 2.1 (Core + SQL), Java, HBase.

- Analyse CRM data:

Creation of a Real Time data pipeline (Lambda Architecture) to analyse sent CRM data, push notifications, SMS, mails and the feedback on this data. Process up to 70 events per seconds.

- Customer 360:

Development of a Customer Recognition application using record linkage. This application uses various sources of data (Real Time Kafka streams / DWH: individual, Booking, Revenue data) to create a more complete customer NoSQL base.

Technical environment: Teradata, TDCH, HBase, Map Reduce, Java, Scala, HDP 2.5, SQL

- Flight delay forecasting:

Prototype a Machine Learning model to forecast Flight Delay using various sources of data (aircraft, passengers, weather, ...). Provide visualization on this predictive model.

Technical environment: HDP 2.3, Spark 1.5 (Core + SQL + ML), R, SQL.

- Various customer scoring projects:

Implementation of customer scoring (as churn score) in a productive environment. Use of PMML file format to import ML models that have been designed by data scientists with R or Python into a new data flow on a Hadoop cluster using Spark Java API. The challenge was to allow Data Scientists to improve their models as often as needed, without any impact on outputs used by business.

Technical environment: Teradata, TDCH, Sqoop, HDP 2.4, HBase, Spark 1.6 (core + SQL), PMML.

Data Science and Big Data Consultant, Sogeti/Cap Gemini, Toulouse, France — 2015-2017

- Data Science leader for our Data Lab:

The Sogeti Analytics Data Lab in Toulouse is used to carry our clients through the current digital shift, and show them how to take full advantage of their data.

Realize innovative projects based on Big Data Architecture with a strong focus on advanced analytics.

Share expertise and skills using a unique collaborative and easy to use platform for all Data Lab projects.

The main success of the Data Lab combined IoT and deep learning, using TensorFlow on top of a Raspberry Pi to ready an equation and give the correct answer talking.

- Pre Sales consultant:

Product presentations and demonstrations. Provide answers to technical questions in a Request for Information (RFI) or a Request for Proposal (RFP). Assist sales in technical and planning aspects of the response.

Data Scientist Intern, French Central Bank, Paris, France — 2015 (6 months)

Implementation of innovative prototypes based on Big Data technologies. Offering a way to demonstrate value added by this project through dashboard creation and automation.

- Proxy Log Data Analysis to detect Cyber Attack or intrusion

Technical Environment: Logstash, Elasticsearch, Kibana, R, Python

- Implementation of a clustering algorithm for anomaly detection in distributed environment.

Technical Environment: Hadoop 2.0 ecosystem (Hortonworks distribution), Spark 1.1

Intern, University of Quebec in Montreal (UQAM), Montreal, Canada — 2014 (3 months)

Research placement on simulation and inference for Stochastic Differential Equations.

- Simulation methods for time-discretised approximation of the trajectory of a process solution of a stochastic differential equation.

- Inference in several SDE models: Vasicek, Geometric Brownian Motion, Cox-Ingersoll-Ross. Research methodology improvement, by finding recent article on this topic. Translating R code to Matlab programs.

Education

National Institute of Applied Science (INSA) Toulouse, France — Engineer/Master Degree in Applied Mathematics, 2015

Mathematics and modelling, Statistics, Machine Learning, Computer Science.

INSA-ISAE, Toulouse, France — Financial Engineering Specialisation, 2015

Financial Engineering. Market Microstructure, Derivative pricing and hedging.

High School Certificate, Montauban, France — Graduated with first class Honours, 2010

Sciences certificate with a specialisation in Mathematics.

Certifications

Professional Scrum Developer (PSD), Toulouse, France — scrum.org, 2016

Scrum (agile methodology) certification with a score of 96%

Test of English for International Communication (TOEIC), Toulouse, France — ETS, 2015

Professional English certification with a score of 895/950

Technical Skills

Programming Languages

- SQL: Good working knowledge of query generation, general logic and coding standards
- Java: Good understanding of object oriented programming, Java is the language I used for delivering final applications, “industrialization”.
- Python: I use mainly Python for Data Science purposes. Specific modules: Pandas, Scikit Learn, TensorFlow and Keras.
- Scala: Ability to write scripts, mainly for quick-win Spark usage.

Databases

- Hadoop ecosystem: Large knowledge of Hortonworks and Cloudera distribution
- Mongo DB and Elasticsearch: Sharding and index selection skills
- HBase: intensive usage
- Teradata: Main relational databases used in my experience

Cloud and micro-services

- Google Cloud Platform
- Docker and Kubernetes basics

Tools & Methodologies

- Atlassian products (JIRA, Bitbucket, Bamboo, Confluence) to improve collaboration
- SonarQube for code quality
- Open alternatives to Atlassian products: GitLab, Trello, Slack
- Scrum developer for a year, then switched to Kanban for another year. At a global level, I have been working using the SAFe (Scale Agility For enterprises) methodology for the last 3 months.

Languages

- French: native
- English: professional proficiency
- Italian and Spanish: beginner

References

- **Aurélie Dupont** - « Data as a Service » dep. Director - Air France KLM
✉ audupont@airfrance.fr ☎ +33 (0)5 62 12 66 81
- **David Gougoud** - « Data as a Service » Big Data Technical Leader – Air France KLM
✉ david.gougoud@gmail.com ☎ +33 (0)6 28 52 28 53
- **Jean-Philippe Arnalot** - Practice Leader Big Data & Analytics - Sogeti / Cap Gemini
✉ jean-philippe.arnalot@sogeti.com ☎ +33 (0)6 63 52 14 00