

# TRISTAN MUSCAT

## DATA SCIENTIST

 Paris, France  Tristan Muscat  +33 6 11 62 52 59  tristan.muscat@protonmail.com

 <https://datartichaut.com>

## EXPERIENCE

### Orange Cyberdefense

*Lead Data Scientist - Sep. 2023 - Oct. 2024, Paris*

**Technical Environment** : Creation of an automatique classification tool for detecting malicious websites fo the Orange Cybersecurite platform.

**Sujet** : Creating a phishing detection model in order to secure online web surfing for all. My scope was :

- MLflow set up as a model registry in order to version and follow models and deploying then as inference servers inside our kubernetes cluster (GKE).
- Airflow tset up to automate data science and engineering related tasks.
- Streamlit as a dashboard to track models, with evidently reports for data drift and Shap for explainability.
- Data pipeline from cleaning to prediction with scikit-learn and imablanced-learn.
- DI/CD pipeline with Gitlab CI.
- Infrastructure management in a kubernetes cluster deployed with Terraform and Helm.

### Kereon Intelligence

*Data Scientist - Nov. 2020 - August. 2022, Paris*

#### Missions

- **Hackathon IMA Innovation, 2022, second place** : "Improvement of a forecasting model based on meteorological data".
- **Projet interne, 2022** : Extraction of text from scans of books in an Azure cloud environment.
- **Ocean Hackathon, 2021** : organisation d'un événement data et innovation en partenariat avec le Village By CA.
- **MACIF, 2021 (SQL, Python)** : Created a database aggregating several data sources to power a real time dashboard.
- **Excelia, 2021** : Worked on the data science part of an online course "Data For Business".
- **Cerfrance, 2021 (Python, Hasura, GraphQL)** : Created an automated data pipeline from extracting data to build a report.
- **France Active, 2020 (Python)** : Created a risk model in a context of bank loans.

## MENTORING

*OPENCLASSROOMS, 2020 - 2024*

Mentor and evaluator for Open Classrooms' training programs : Data Scientist, Data Analyst and Ingénieur Machine Learning.

## SKILLS

### Data Science

Programming : Python, R.  
DS Libraries : scikit-learn, H2O.  
Softwares : Dataiku.

### Data Engineering

DBMS : SQL, MongoDB, Redis.  
Backend : Celery, FastAPI.  
Tools : MLflow, Airflow.

### Data Visualisation

Python : Dash, Streamlit.  
Softwares : Tableau, PowerBI, Dataiku.

### Data Management

RGPD.  
Data Quality : Great Expectation.

### Cloud

Azure Certified.  
GCP : one year experience.  
AWS three year mentoring.

### Build

Docker.

### Deployment

Kubernetes / Helm / Terraform.

### CI/CD

Gitlab CI.  
Argo CD.

## LANGUAGES

- **French** (native speaker)
- **English** (fluent)

## EXPERIENCE

### AddixWare

*Data Scientist* - Nov. 2019 - Nov. 2020, Paris

#### Missions

- **LVMH, 2020 (Python)** : modeling of a physical value (flash point) for cosmetics in order to comply with requirement specifications.
- **Covid-19, 2020 (Python)** : Study about comorbidity regarding covid-19 in order to give facts to base published articles off of. Data were gathered from a Chinese and Mexican hospital giving diagnosis and personal information.
- **Sanofi Pasteur, 2019 (Python)** : Data science support to create a patented vaccine production process.

### Magellan Consulting

*Data Scientist* - May 2019 - August 2019, Paris

#### Missions

- **Crédit Agricole, 2019** : Travaux de cadrage de refonte de la gestion des connaissances à l'échelle Groupe.
- **Interne, 2019 (Python, TensorFlow, Android Studio)** : Travaux sur le thème des services cognitifs en machine learning: computer vision (Convolutional Neural Network).

### Voirin Consultants

*Data Scientist* - April 2018 - January 2019, Paris

#### Missions

- **SIAAP, 2018** : Writing of a data management document as a vademecum, about data usages for enterprises, from implementation data governance to launching solutions in production.
- **Caen-la-mer, 2018** : Élaboration d'un schéma directeur des systèmes d'informations pour une communauté urbaine.

### Optimind Winter

*Data Scientist* - March 2016 - April 2018, Paris

#### Missions

- **Carte Blanche, 2018** : Cost analysis of the «Reste à Charge Zéro» health project.
- **Prévère, 2017, RPA (UiPath)** : RPA (UiPath). Developed a program with UiPath allowed to automatically process PDF files.
- **La Poste, 2017 (R, RShiny)** : Absenteeism modeling and analysis with data and machine learning techniques.
- **EURIA, 2017, (R)** : Helped and accompanied data science study groups on two subjects.
- **Interne, 2017, (R)** : Credit rating migration modeling in order to optimize the Solvency Capital Requirement (SCR).
- **Barclay's, 2017, (R)** : liability cash flow modeling with machine learning (deposit and savings accounts).
- **Predica, 2016, (R)** : Customer behavior modeling in life insurance (savings account).

## EDUCATION

### Microsoft Azure Certifications

2023, Paris

AZ-900 - DP-900 - AI-900

### Dataiku Certifications

2023, Paris

Dataiku Core Designer

### Université Dauphine, Paris

2014-2016, Paris

M1 MIDO, M2 Actuariat

### E.I.S.T.I.

2013-2016, Paris

Engineering Degree

### Université de Cergy Pontoise

2013-2014, Paris

Licence Informatique

### Lycée Saint Joseph, Avignon

2011-2013, Avignon

Classe Préparatoire (CPGE-PC)

### Lycée Saint Louis, Orange

2012, Orange

Baccalauréat Scientifique

## INTERESTS

### Kactu'Z Actuariat, 2017 -

planning for the "journée Kactu'Z de l'actuariat" and the "Nuit Des Actuaire's".

### LADD Treasurer, 2018 -

Planning and organizing events for the Dauphine's Actuarial Sciences Master students.