

Christian NGUIEPE SEGNOU

Data Scientist, AI Engineer, Formateur Data
L’Haÿ-les-Roses, Île-de-France

Phone : 0666484774
Email : nguiepemarius@gmail.com
LinkedIn : <https://www.linkedin.com/in/christian-segnou-ph-d-4461b2102/>
Github : <https://github.com/ChrisDMI>
Portfolio : <https://chrisdmi.github.io/Portfolio-christian-segnou-2023/>
CV approche RAG : <https://christian-segnou-cv.streamlit.app/>

Introduction

Expert en computer vision et en traitement d'image, j'ai acquis une vaste expérience en analyse de données et en machine learning durant ma thèse. Formateur en machine learning, j'aide les salariés à maîtriser les compétences en data science nécessaires pour réussir. Mon parcours en intelligence artificielle et en data science se distingue par des projets innovants et une capacité à résoudre des problèmes complexes en utilisant des technologies de pointe. Mon expertise couvre également le développement et l'intégration de modèles d'apprentissage automatique dans des environnements industriels. Passionné et curieux, je suis toujours à la recherche de nouveaux défis pour apporter des solutions efficaces et créatives.

Compétences & Environnements techniques

Data Science, Data Analyse, Data Engineering

Python, PySpark, SQL, Git CI/CD, Docker, MongoDB, Kubernetes, KubeFlow, MLFlow, Airflow, Linux, HTML/CSS, Scrapy, Streamlit, FastApi, Pytorch, Tensorflow, AWS

Formation & Diplômes

Année	Diplôme	Ecole	Spécialité
2023	Doctorat	Université de Paris-Saclay	Sciences des Matériaux
2023	Master 2	Jedha	Concepteur Developpeur en Science des Données
Langues	Français (Langue Maternelle), Anglais (courant)		

Certifications

- **Concepteur Developpeur en Science des Données** - <https://www.credential.net/8ae1ca89-eab2-44a6-9bed-98fe27562a77#gs.5fw763>
- **Hands-on to Linux Commands and Shell Scripting** - <https://coursera.org/verify/AUH4LGP2NCT4>
- **Relational Databases (RDBMS) Essentials** - https://www.credly.com/badges/4ae1e959-dbfd-4518-9178-73cea91b9edf/linked_in_profile
- **Introduction to Relational Databases (RDBMS)** - <https://coursera.org/verify/2B8WSU3F7V74>
- **Relational Database Administration (DBA)** - <https://coursera.org/verify/MSBMTEKGBTHC>
- **Databases and SQL for Data Science with Python** - <https://coursera.org/verify/8EGZSDB8QRK3>
- **Data Engineering Essentials** - https://www.credly.com/badges/2c1c40e5-8e58-4cbd-bb61-fc0cf2d975ee/linked_in_profile
- **Introduction to Data Engineering** - <https://coursera.org/verify/WKEGFMRJXYBK>
- **Python for Data Science, AI & Development** - <https://www.coursera.org/account/accomplishments/verify/FFK2VE8FG4QN>
- **Python Project for Data Engineering** - <https://coursera.org/verify/BG8EXTHBUBSP>
- **ETL and Data Pipelines with Shell, Airflow and Kafka** - <https://coursera.org/verify/ZWPY485NAFQH>
- **HTML, CSS, and Javascript for Web Developers** - <https://www.coursera.org/account/accomplishments/verify/3MFJVEVM8SX7>

Expériences Professionnelles

03/2024 –	Formateur Freelance chez Abilycare
Fonction	Formateur en Machine Learning
Projet	Formation en compétences data pour employés non spécialistes
Responsabilités	<ul style="list-style-type: none">- Préparation et animation de cours théoriques et pratiques, de travaux pratiques (TP) et exercices en statistiques appliquées à la data science.- Enseignement des techniques de prétraitement des données pour optimiser leur analyse et leur exploitation.- Formation au machine learning supervisé et non supervisé, permettant une compréhension approfondie des modèles prédictifs de regression et de classification.- Initiation aux fondamentaux du deep learning, avec des applications concrètes en vision par ordinateur (CV) et en traitement automatique du langage naturel (NLP).
Environnement Technique	Python, Jupyter Notebook, Discord, Vulgarisation, Pédagogie

09/2023 –	Expleo Group – Montigny-le-Bretonneux
Fonction	Data Scientist R&D
Projet	Repair Smart : Extraction et Analyse des informations dans les documents PDF d’AirBus
Responsabilités	<ul style="list-style-type: none">- Construction d’un modèle YOLO pour la détection des zones d’intérêt dans les fichiers PDF- Refactoring de code pour faciliter l’utilisation du modèle FasterRCNN- Pré-traitement des crops et application d’un modèle OCR pour la reconnaissance de texte- Construction d’une application pour mettre à disposition le modèle.- Développement et intégration d'un chatbot, connecté à un Large Language Model (LLM), pour fournir un système de questions-réponses dynamique basé sur le contenu des documents PDF.- Implémentation d'un Retrieval-Augmented Generation (RAG) avec Langchain pour améliorer la précision et la pertinence des réponses, facilitant ainsi l'accès aux informations cruciales pour les utilisateurs
Challenges	<ul style="list-style-type: none">- Gestion de la diversité des patterns dans les documents, nécessitant l'élaboration de méthodes d'analyse flexibles capables de s'adapter à divers formats et structures de données.- Défi de la reconnaissance d'informations écrites manuellement, impliquant l'implémentation et l'optimisation d'outils de reconnaissance optique de caractères (OCR) avancés pour assurer une extraction précise des données manuscrites.
Environnement Technique	Python, PyTorch, YOLO, FasterRCNN opencv, Pytesseract, Kubeflow, Git, LLM, Langchain

Fonction	Data Scientist R&D
Projet	Automatisation et Optimisation, à l’aide de l’IA, de la Gestion d’Équipements pour le Département Logistique
Responsabilités	<ul style="list-style-type: none">- Construction d’un modèle YOLO pour la détection des zones d’intérêt dans les fichiers PDF- Pré-traitement des crops d’images en vue d’une reconnaissance de texte- Application d’un modèle OCR pour la reconnaissance de texte- Construction d’une application pour mettre à disposition le modèle.- Construction d’une base de données et remplissage de cette base après chaque analyse de fichier pdf.- Construction et déploiement d’une application pour mettre à disposition le modèle.
Challenges	<ul style="list-style-type: none">- Assurance de l'exactitude des informations extraites, essentielle pour l'intégrité de l'inventaire des équipements et le respect des procédures de restitution.- Mise en œuvre rapide d'une solution automatisée intégrée aux processus existants, répondant aux exigences urgentes du département logistique.
Environnement Technique	Python, PyTorch, YOLO, opencv, Pytesseract, Kubeflow, Git, Docker, Streamlit, déploiement

09/2019 – 09/2021, 2 ans	Université de Paris-Saclay - Orsay
Fonction	Enseignant de Python
Projet	En charge de la préparation et de la dispense, aux étudiants des niveaux L1 a L3, des cours sur la programmation Python et sur les concepts d'introduction du Machine learning et du Deep learning.
Responsabilités	<ul style="list-style-type: none"> - Développement et animation de cours interactifs - Accompagnement des étudiants dans leurs projets de fin de cours - Evaluation des progrès des étudiants et fourniture de retours constructifs pour améliorer leur compréhension
Environnement Technique	Python, Vulgarisation, Pédagogie

10/2018 – 06/2023, 5 ans	Centre de Nanosciences et de Nanotechnologie - Palaiseau
Fonction	Phd Candidate
Projet	Project Lead – Comprendre la dynamique des parois de domaine magnétique pour l'amélioration des disques durs
Responsabilités	<ul style="list-style-type: none"> - Traitement d'image issus de la microscopie magnéto-optique Kerr - Analyse des données issus de la mesure du déplacement des parois de domaine magnétique - Modélisation mathématique de la dynamique d'une paroi de domaine - Prediction à partir du machine learning des vitesses des parois
Environnement Technique	Python, Scikit-Learn, Pandas, opencv, ImageJ, Git, Numpy, Scipy, Matplotlib, Sympy