

Nathan Randriamanana

Administrateur de répertoire - Data Scientist - Statisticien | né le 28/07/1999

Meudon, 92190, France

+33 6 24 05 82 28 | nathan.randriamanana@insee.fr | GitHub | LinkedIn



Parcours professionnel

Data Scientist - MOA du répertoire Sirene

Direction des statistiques d'entreprises - DG Insee

Montrouge

Début Septembre 2023

- Pilotage de l'intégration de la NAF 2025 en flux pour la double codification de l'APE dans l'application Sirene 4 de la conception des scénarios à la mise en œuvre agile avec un MVP livré avant échéance
- Réentraînement des modèles de Machine Learning pour la codification de l'APE en double nomenclature intégrant la NAF actuelle et la NAF 2025
- Conduite de deux campagnes d'annotations majeures pour les réentraînements des modèles de codification incluant le déploiement technique de la plateforme et l'animation du réseau d'experts métier
- Modernisation de l'architecture de codification via une approche MLOps garantissant la reproductibilité et la maintenabilité des modèles
- Référent métier et appui transversal pour la gestion des incidents de production incluant la résolution des cas complexes au répertoire et le support aux gestionnaires utilisateurs
- Interface entre les équipes métier et informatique via la rédaction de spécifications fonctionnelles et le suivi de leur implémentation par la MOE
- Conception d'outils de pilotage et d'explicabilité incluant des tableaux de bord pour le suivi du taux d'automatisation et des dispositifs d'aide à la décision en self-service
- Contribution aux groupes de veille technologique dédiés aux innovations en intelligence artificielle
- Animation d'ateliers internes sur les bonnes pratiques de développement en self pour favoriser l'autonomie du métier

Stagiaire Data Scientist (fin d'étude)

Data Science Hub - Statistic Flanders

Bruxelles, Belgique

Mars 2023 - Août 2023

- Contextualisation via des méthodes de NLP de l'indicateur de sentiment de la population flamande obtenue par l'analyse des sentiments sur les tweets du réseau social Twitter/X
- Analyse de séries temporelles pour caractériser la saisonnalité de la satisfaction de la population
- Usage de scénarios de topic modeling pour catégoriser les événements vécus quotidiennement par la population
- Exploration d'indicateurs statistiques pour détecter les événements contribuant à des pics de satisfaction

Stagiaire DevOps

Division Innovation instruction technique - DG Insee

Montrouge

Juin 2022 - Juillet 2022

- Conception de formations sur la plateforme "SSP Cloud", instance d'Onyxia hébergée à l'Insee
- Montée en compétence sur les technologies cloud, culture de l'open source et de la data

Stagiaire Data Scientist sur la cybercriminalité

Service Statistique du Ministère de l'Intérieur (SSMSI)

Bercy

Mai 2021 - Juillet 2021

- Détection des nouvelles manières d'opérer sur la cybercriminalité
- Statistiques sur la délinquance, cybercriminalité, machine learning, NLP

Expertise et Engagements

Formation des enseignants de l'Ensea au MLOps

Formation GENES - Projet PHAS de la Banque Mondiale

Abidjan, Côte d'Ivoire

Octobre 2025

Participation à l'encadrement de la seconde édition de Datafid

Défi Datafid - Projet Data Governance in Africa - Expertise France, financé par l'UE

Abidjan, Côte d'Ivoire

Mars 2025

Encadrement d'un hackaton sur les thématiques fiscales et douanières pour l'Afrique

Défi Datafid - Projet Datafid - Expertise France, financé par la DG Trésor

Abidjan, Côte d'Ivoire

Juin 2023

Participation au programme 10% (Saison 2)

Projet « Transformers, une IA open-source pour l'analyse de texte par les administrations ».

Paris
Septembre 2023 -
Novembre 2023

Dataiku Challenge

Mise au point d'une application pour extraire les insights (résumé, analyse des sentiments, ...) d'un article de journal en utilisant la plateforme Dataiku

Ensaï, Bruz
Février 2022

Tech Sprint de l'ACPR

Hackathon organisé par le Pôle Fintech-Innovation de l'Autorité de contrôle prudentiel et de résolution (ACPR) sur l'explicabilité des algorithmes de risque de crédit en "boîte noire".

Banque de France, Paris
Juillet 2021

Compétences

Soft skills

Savoir-être

- Bon sens du service public
- Gestion des priorités
- Travail en équipe
- Capacité de synthèse
- Clarté
- Créativité
- Respect des échéances

Programmation

R, Python, SQL, Git (Github/Gitlab), Docker, Kubernetes, Bash, Shell

Balisage

Markdown, Typst, Quarto, LaTeX

Langages parlés

Anglais, Allemand (notions), Malgache

Formations

Master Science des Données pour la Décision Publique

Majeure Méthodologie de la statistique publique (MSP), labellisé EMOS (European Master in Official Statistics)

Bruz, France
2022 - 2023

CPGE (classe préparatoire aux grandes écoles) scientifique

Filière MPSI-MP (mathématiques, physique et sciences de l'ingénieur)

Reims, France
2017 - 2020

Baccalauréat général scientifique (Mention Très Bien)

Option Sciences de l'Ingénieur, Spécialité ISN (Informatique et Sciences du Numérique), Section européenne Anglais

Pontault-Combault,
France
2014 - 2017

Communications

Contributions aux Journées de méthodologie statistique de l'Insee

2 soumissions sur le sujet de la codification dont

Jussieu
Novembre 2025

- l'utilisation de LLM pour entraîner un classifieur en nouvelle nomenclature
- la présentation de la nouvelle librairie TorchTextClassifiers

Diverses présentations auprès de divers interlocuteurs sur le sujet de la codification

Présentations et échanges notamment auprès de l'Inserm, Statistics Netherlands, Statec Luxembourg, EuroStat, l'UNSD et le SSP hub, le réseau des data scientists du Service Statistique Publique

webinaires en ligne