

# VINCENT ROCHER

Recherche d'emploi en data science pour la fin de l'année 2021.

À la recherche d'un emploi dans le domaine de la **Data Science - Intelligence Artificielle**, pour commencer à la fin de mon doctorat qui se termine en fin d'année 2021.



## EXPÉRIENCE PROFESSIONNELLE

2018  
|  
Présent

### ● Doctorant en Bioinformatique

Centre de Biologie Intégrative (CBI) - Équipe Chromatin and DNA repair

📍 Toulouse, France

Pendant ma thèse, j'ai travaillé sur des projets à long terme en collaboration avec des chercheurs en biologie, statistique et informatique provenant de plusieurs institutions.

#### Compétences générales :

- Présenter & Rédiger des résultats de recherche pour des conférences & articles scientifiques.
- Veille bibliographique pour déterminer les méthodes statistiques adéquate aux données biologiques.
- Écrire les méthodes statistiques utilisées pour publication.
- Vulgarisation scientifique pour les chercheurs & les étudiants.
- Enseignement en Master de méthodes statistiques générales & spécialisés pour les données biologiques.

#### Traitement, Analyse, Inférence et Visualisation de données biologiques. :

- Traitement & Analyse des données biologiques "brutes" de façon à les rendre exploitables via différentes pipelines informatiques.
- Expliquer & visualiser des données complexes en utilisant des méthodes simples.
- Inférence des données biologiques en utilisant des méthodes statistiques paramétriques ou non-paramétriques.
- Mise au point de modèles de Machine/Deep Learning pour l'explication & la prédiction de données biologiques.
- Réduction de dimension pour expliquer des jeux de données à haute dimension.

2015  
|  
2018

### ● Apprenti Ingénieur - Ingénieur d'Étude en bioinformatique

Centre de Biologie Intégrative (CBI) - Équipe Chromatin and DNA repair

📍 Toulouse, France

- Mise au point de l'infrastructure logicielle pour le traitement des données de l'équipe - laboratoire sur serveur de calcul.
- Mise au point des pipelines d'outils bioinformatiques pour le traitement des données biologiques.
- Automatisation du traitement des données biologiques.
- Mise en place d'applications web "user-friendly" pour le traitement de données biologiques.
- Collaboration avec le biologiste pour la mise au point du plan d'expérience.
- Écriture du matériels & méthodes - Publication des scripts sur GitHub.

## CONTACT

- ✉ [vincent.rocher@univ-tlse3.fr](mailto:vincent.rocher@univ-tlse3.fr)
- ✉ [vincent.rocher.armd@gmail.com](mailto:vincent.rocher.armd@gmail.com)
- 🌐 [github.com/rochevin](https://github.com/rochevin)
- 📞 06 70 78 40 84

Pour plus d'information, vous pouvez me contacter par mail.

## PROGRAMMATION

Programmation en **R** & **Python** (bases en C).  
Ligne de commande avancée avec **bash**.  
**Git** & gestion de versions.

## DATA-SCIENCE

Automatisation de pipelines avec **Snakemake** & **Luigi**.  
Manipulation, Traitement & Visualisation des données avec **R** (**tidyverse/ggplot**) & **Python** (**numpy/pandas**).  
Inférence et prédiction en utilisant des modèles de **Machine Learning** paramétriques & non-paramétriques avec **tidymodels** & **scikit-Learn**.  
Construction de modèles de **Deep Learning** avec **Keras** & **Tensorflow**.

## LANGUES

Anglais scientifique : **890** au TOEIC.

Document généré le : 2021-06-20.



## PUBLICATIONS SÉLECTIONNÉES

2021

### • **DeepG4: A deep learning approach to predict active G-quadruplexes from DNA**

Preprint sur bioRxiv - Soumission à Plos Computational Biology.

[Vincent Rocher](#), Matthieu Genais, Elissar Nassereddine and Raphael Mourad.

- Mise au point du modèle de Deep Learning & production des graphiques de la publication.
- Développement d'un package R pour englober l'outil.

2021

### • **Loop extrusion as a mechanism for formation of DNA damage repair foci**

Nature 2021.

Coline Arnould, [Vincent Rocher](#),... and Gaëlle Legube.

- Traitement & analyse des données biologiques publiées.
- Manipuler des données de très grandes dimensions sous la forme de **sparse matrix** et simplifier ces données de façon à produire un résultat compréhensible & publiable.
- Intégrer des données multi-variées.

2018

### • **Comprehensive Mapping of Histone Modifications at DNA Double-Strand Breaks Deciphers Repair Pathway Chromatin Signatures**

Mol Cell. 2018.

Thomas Clouaire, [Vincent Rocher](#),... and Gaëlle Legube.

- Traitement et Analyse des données biologiques publiées.
- Production des figures et inférences statistiques sur la population.



## ENSEIGNEMENT

2019  
(15h)

### • **Introduction aux statistiques**

Enseignement TP pour les Master 1 en biologie.

📍 Université de Toulouse, France

2020  
(20h)

### • **Bioinformatique & Traitement de données de séquençage à haut débit**

Enseignement TP pour les Master 1 Bioinformatique.

📍 Université de Toulouse, France



## ÉDUCATION

2018  
|  
Présent

### • **Doctorat en Bioinformatique**

Université Toulouse 3 Paul Sabatier

📍 Toulouse, France

2014  
|  
2017

### • **Master en Bioinformatique en apprentissage en alternance (format spécial en 3 ans)**

Université de Rouen

📍 Rouen, France