

# ANTONIN POCHÉ

✉ poche.antonin@gmail.com ☎ +336 31 57 27 34  
📍 8 Avenue de l'aérodrome de Montaudran, 31400 Toulouse, France



## PhD Student in Interpretability for NLP

### SHORT INTRODUCTION

I am a second-year PhD student in interpretability for language. Core maintainer of the Interpreto open-source library. My goal is to provide easy access to useful explanations.

### FORMATION

#### PhD Student in XAI for NLP

📅 October 2024 - On going  
📍 IRT Saint Exupéry & IRIT - Toulouse



Subject: Paving the way toward conversational explainability.

Direction: Pr. Philippe Muller, Pr. Nicholas Asher, and Dr. Fanny Jourdan.

#### UGA Grenoble INP Ensimag

📅 Sept 2017 - June 2021  
📍 Grenoble



Master: Applied Mathematics and Computer Science

#### Politecnico di Torino

📅 Sept 2019 - June 2021  
📍 Turin



Master: Data Science and Computer Engineering

### WORK EXPERIENCE

#### Research Engineer in Explainability

📅 June 2021 – October 2024  
📍 IRT Saint-Exupéry - Toulouse



Applied research in XAI (eXplainable Artificial Intelligence) on two projects for trustworthy AI: DEEL.ai and Confiance.ai.

#### Research Intern in XAI for RL

📅 Nov 2020 – May 2021  
📍 Bordeaux



Applied research internship on explainable deep reinforcement learning.

#### Bio-informatics Research Intern

📅 June 2019 – Aug 2019  
📍 Laboratoire Jean Kuntzmann - Grenoble



Research internship in bio-informatics on deep learning for genomic data.

### HOBBIES

⚽ Sports

🌲 Nature

🎮 Games

### TEACHING

#### Mastere AIBT

📅 October 2021 - On going  
📍 ISAE Supaero



AIBT: Artificial Intelligence and Business Transfer.  
Practical sessions on AI and lessons on explainability.

#### Mastere ValDom

📅 April 2021 - On going  
📍 Enseeeht  
[0.5cm] ValDom: "Valorisation des Données massive".  
Lessons on explainability.



### SCIENTIFIC CONTRIBUTIONS

#### ConSim: Measuring Concept-Based Explanations' Effectiveness with Automated Simulatability

📅 July 2025  
📍 ACL 2025



#### Interpreto: Interpretability Toolbox for LLMs

📅 June 2025  
📍 GitHub

#### Natural Example-based Explainability: A Survey

📅 June 2023  
📍 World Conference on xAI



#### Xplique: A Deep Learning Explainability Toolbox

📅 June 2022  
📍 CVPR - XAI4CV Workshop  
Open Source XAI library with over 650 GitHub stars.



#### Guidelines to Explain ML Algorithms

📅 September 2023  
📍 DEEL

### LANGUAGES

French

Native

English

Advanced

Italian

Intermediate