Jean-Charles **Noirot Ferrand**

M.S. Student · Computer Sciences

1210 W. Dayton Street, Madison, WI 53706-1613, USA

jcnf@cs.wisc.edu https://jcnf.me Jean-Charles Noirot Ferrand

Research Experience

MadS&P - Security and Privacy Research Group at UW–Madison — *UW–Madison*

Research Assistant — *2023 - Present*

- My research focuses on understanding the robustness of large language models in adversarial settings

LIRIS Laboratory, École Centrale de Lyon — *Centrale Lyon*

Research Internship — *Summer 2023*

- Development of DESCRIPT, a tool for creating and visualizing stroke-based drawings

Education

University of Wisconsin–Madison — *Madison, WI, USA*

M.S. in Computer Sciences — *2023 - Present*

- Research at the intersection of adversarial machine learning and large language models.
- Advisor: Prof. Patrick McDaniel

École Centrale de Lyon — *Lyon, France*

Diplôme d'Ingénieur (M.S. and B.S. in Engineering Sciences) — *2021 - Present*

- French Engineering School
- Multidisciplinary studies: Maths, Physics, Computer Science, Fluid Mechanics, Electrical Engineering, Economics, Management, etc.

Claude Bernard Lyon 1 University — *Lyon, France*

B.S. in General Mathematics and Applications — *2021 - 2022*

- Dual degree during the first year at École Centrale de Lyon

Lycée Marcelin Berthelot — *Saint-Maur-des-Fossés, France*

Classe Préparatoire aux Grandes Écoles (equivalent to the first 2 years of B.S. in Engineering Sciences) — *2019 - 2021*

- 2-year intensive program preparing for the national competitive exams for entry to the top French Engineering Schools

Publications

Conferences

- Kunyang Li, Kyle D Domico, **Jean-Charles Noirot Ferrand**, and Patrick McDaniel. "The Efficacy of Transformer-Based Adversarial Attacks in Security Domains". In: *MILCOM 2023 - Workshop on Artificial Intelligence for Cyber (MILCOM 2023 - Workshop on AI for Cyber)*. Boston, USA, Oct. 2023, p. 6.

Thesis

- Jean-Charles Noirot Ferrand. "Extracting the Harmfulness Classifier from Aligned LLMs". University of Wisconsin-Madison, Dec. 2024.

Professional Activities

Reviewer - Conferences

2025 **International Conference on Learning Representations (ICLR),** External Reviewer

2025 **USENIX Security Symposium (USENIX Security),** External Reviewer

2025 **IEEE Symposium on Security and Privacy (IEEE S&P),** External reviewer

2024 **ACM Conference on Computer and Communications Security (ACM CCS),** External Reviewer

Other Service

2024 **UW-Madison New CS Graduate Students Mentoring Program**, Mentor

Skills

Computing Skills

Languages C, C#, Bash, Python, Matlab, LaTeX, HTML, CSS, JavaScript, SQL
OS GNU/Linux (Ubuntu), Windows
Softwares Docker, Git, HTCondor

Language Skills

French Native
English Fluent
Spanish Intermediate
Chinese Beginner